Energy-Efficient Delay-Tolerant Cognitive Radio Networks

Zhao, Bi

Awarding institution:
King's College London

King's College London

Energy-Efficient Delay-Tolerant Cognitive Radio Networks

by

Bi Zhao

Submitted to the King's College London, for the degree of Doctor of Philosophy

Centre for Telecommunications Research
Department of Informatics

July 2015

"A man can succeed at almost anything for which he has unlimited enthusiasm"

Charles M. Schwab

Abstract

Undoubtedly the upward growth trend in aggregate mobile Internet IP traffic is expected to continue steadily. In this emerging mobile environment with increased data traffic and always-on applications, the limitations of battery technologies lead to drastically shorten recharging cycles of mobile devices. For mobile applications that can tolerate a moderate delay which account for a high proportion of global mobile traffic, a technique that postpones the data transmission to high-rate hotspots could effectively provide significant energy gains, which can be translated to increased battery lifetime.

In this thesis, potential gains are explored resulting from utilization of two important technologies for future and emerging wireless networks, namely Cognitive Radios and Delay Tolerant Networking. Both of them are in essence opportunistic in their operation and so far have been considered in isolation. Considering that an increased number of worldwide countries are permitting operation of cognitive radio systems in the spatially vacant licensed analog TV bands, this would enable new possibilities to provision further capacity increase for wireless broadband and multimedia services. Hierarchical CR networks improve spectrum efficiency by allowing the low-priority SUs to temporarily seek the wireless spectrum that is licensed to different organizations. Once mobile devices are equipped with multiple air-interfaces allowing them to connect to cellular networks, Wi-Fi and White-Fi, they could switch among these networks to seek and use any licensed spectrum bands as long as they avoid interference being caused to TV receivers. When wireless nodes are competing for secondary access to the medium, the estimation of probability of PU arrival rate and service time is important for mobile devices (SU) to effectively occupy the primary spectrum. The mobile nodes firstly contact a trusted database for historical information about PU traffic at a specific location and time duration so as to estimate the probability for the SU connections. Then, regarding the SU traffic, it is shown that it can be modeled as an M/M/K/L queuing system which allows to analyze the capability that the system can serve users simultaneously.

As location of mobile users is the key to determine the capacity of accessible wireless service for themselves, stochastic characteristics of user mobility are studied in terms of user velocity, direction changes, and route selection distribution. Moreover, when mobile terminals are moving among different cells supporting different network technologies, the

performance of vertical handover and cell residence time in the coverage of Wi-Fi/White-Fi hotspots would greatly affect the overall efficiency of wireless transmission. In this scenario, if the mobile applications could tolerate some delay, the proposed schemes can significantly avoid the drain of mobile device batteries by making selective use of the nearby high-speed hotspots.

Nowadays, with the surge of the diverse and ubiquitous Internet applications, mobile users expect to enjoy wireless Internet connectivity anywhere and at any time. According to the inherent mobility of mobile users, optimal stopping problem is formulated for energy-delay trade-off, which the stopping decision would be made based on channel conditions, delay constraints, and energy cost. In addition, for popular video streaming applications on portable devices that could be watched several times by one user, there are trade-offs between storing video content locally at the DRAM of the device or allowing deleting the content from the local memory and relaying in wireless streaming in near-future requests of the same content. To this end, a scheme has been proposed where the mobility of the user is taken into account together with the probability of a user requesting the same content multiple times so that a decision is taken of whether or not the content should be stored locally.

Finally, since the proliferation of always-on Internet applications has put significant s-train on the battery capabilities, the problem of prolonging battery lifetime of mobile devices is introduced. Previous research has revealed that the data downloading via wireless radios is a dominant energy consumption factor in mobile devices. To avoid the drain of mobile device batteries, based on the delay tolerance of mobile Internet applications, the proposed strategies are designed in which mobile terminals can intelligently switch among cellular, Wi-Fi and White-Fi interfaces, by considering the energy cost, RF coverage, capabilities, and transmission algorithm. Numerical experiments on various mobility models reveal that the energy cost of wireless transmission closely relates to user locations, mobility pattern, spectrum availabilities as well as applications' delay tolerance and available wireless access technologies.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Dr Vasilis Friderikos for his guidance through many interesting discussions, brainstorming sessions, and assistance whenever needed.

Further, I would like to thank Professor Hamid Aghvami and my colleagues in CTR, who have given me valuable perspectives on my work, and other friends at King's College London, who have contributed to a friendly and joyful working environment.

Finally, I would like to express my great appreciation to my parents and my wife for their endless love and support, which is the main driver for me to pursue my PhD. They made me who I am.

Contents

List of Figures

List of Tables

Abbreviations

ADC	**A**nalog-to-**D**igital **C**onverter
AMC	**A**daptive **M**odulation and **C**oding
AP	**A**ccess **P**oint
API	**A**pplication **P**rogramming **I**nterface
BP	**B**undle **P**rotocol
BS	**B**ase **S**tation
BU	**B**inding **U**pdate
CDF	**C**umulative **D**istribution **F**unction
CDMA	**C**ode **D**ivision **M**ultiple **A**ccess
CN	**C**orrespondent **N**ode
CoA	**C**are-**o**f **A**ddress
CPU	**C**entral **P**rocessing **U**nit
CR	**C**ognitive **R**adio
CRT	**C**ell **R**esidence **T**ime
DAG	**D**ata **A**cquisition and **G**eneration
DLL	**D**elay-**L**ocked **L**oop
DRAM	**D**ynamic **R**andom-**A**ccess **M**emory
DSA	**D**ynamic **S**pectrum **A**llocation
DSL	**D**ynamic **S**trategy **L**earning
DTN	**D**elay **T**olerant **N**etwork
D2D	**D**evice **T**o **D**evice
ECC	**E**lectronic **C**ommunications **C**ommittee
EIS	**E**nhanced **I**nformation **S**erver
FCFS	**F**irst **C**ome **F**irst **S**erve
FIFO	**F**irst **I**n **F**irst **O**ut

FTP	**F**ile **T**ransfer **P**rotocol
GPRS	**G**eneral **P**acket for **R**Radio **S**ervice
GSM	**G**lobal **S**ystem for **M**obile Communications
GSPI	**G**eneric **S**erial **P**eripheral **I**nterface
HA	**H**ome **A**gent
HC	**H**andover **C**oordinator
HDD	**H**ar**D** **D**isk
HoA	**Ho**me **A**ddress
HSDPA	**H**igh-**S**peed **D**ownlink **P**acket **A**ccess
HTTP	**H**yper**t**ext **T**ransfer **P**rotocol
IETF	**I**nternet **E**ngineering **T**ask **F**orce
i.i.d.	**i**ndependent and **i**dentically **d**istributed
IP	**I**nternet **P**rotocol
LCD	**L**iquid **C**rystal **D**isplay
LEO	**L**ow **E**arth **O**rbit
LMA	**L**ocal **M**obility **A**nchor
LTE-A	**L**ong **T**erm **E**volution **A**dvanced
LTE	**L**ong **T**erm **E**volution
LTP	**L**icklider **T**ransmission **P**rotocol
MAC	**M**edia **A**ccess **C**ontrol
MAG	**M**obile **A**ccess **G**ateway
MIPv6	**M**obile **IPv6**
MN	**M**obile **N**ode
NAND	**N**egated **AND**
OS	**O**perating **S**ystem
OSP	**O**ptimal **S**stopping **P**rogramming
PC	**P**ersonal **C**omputer
PCI	**P**eripheral **C**omponent **I**nterconnect
PCMCIA	**P**ersonal **C**omputer **M**emory **C**ard **I**nternational **A**ssociation
PDA	**P**ersonal **D**igital **A**ssistant
PMIPv6	**P**roxy **M**obile **IPv6**
PoA	**P**oint **o**f **A**ttachment
PRP	**P**reemptive **R**esume **P**riority

PS	**P**rocessor **S**haring
PU	**P**rimary **U**ser
P2P	**P**eer **T**o **P**eer
QAM	**Q**uadrature **A**mplitude **M**odulation
QoS	**Q**uality **o**f **S**ervice
QPSK	**Q**uadrature **P**hase-**S**hift **K**eying
RF	**R**adio **F**requency
RSS	**R**ich **S**ite **S**ummary
RSSI	**R**eceived **S**ignal **S**trength **I**ndicator
RVS	**R**andom **V**ariable**S**
RWP	**R**andom **W**ay**p**oint
SD	**S**ecure **D**igital
SDIO	**S**ecure **D**igital **I**nput **O**utput
SU	**S**econdary **U**ser
TCP	**T**ransmission **C**ontrol **P**rotocol
TD-SCDMA	**T**ime **D**ivision-**S**ynchronous **C**ode **D**ivision **M**ultiple **A**ccess
TTI	**T**ransmission **T**ime **U**nit
TV	**T**ele**v**ision
TVWS	**TV W**hite **S**pace
UDP	**U**ser **D**atagram **P**rotocol
UE	**U**ser textbfEquipment
UGC	**U**ser-**G**enerated **C**ontent
UMTS	**U**niversal **M**obile **T**elecommunications **S**ystem
VHDA	**V**ertical **H**andoff **D**ecision **A**lgorithm
WiMAX	**W**orldwide **i**nteroperability for **M**icrowave **A**ccess
WLAN	**W**ireless **L**ocal **A**rea **N**etwork
3GPP	**3**rd **G**eneration **P**artnership **P**roject

Chapter 1

Introduction

Over the last few years we are witnessing a significant increase in the aggregate traffic in mobile networks which is due to the proliferation of smartphones and mobile Internet applications. In this environment mobile users expect to enjoy ubiquitous wireless Internet experience which boils down to providing high capacity connectivity to them anywhere and at any time. For sustainability reasons operational as well as capital expenditure for mobile operators will need to be reduced. Energy consumption plays a significant role in the overall operational expenditure of a mobile operator. To this end, significant efforts have been recently placed on reducing the overall energy consumption leading to the so-called green networks.

Significant research efforts have been placed on energy efficient data transmission that takes into account the limited energy resources and/or the delay-sensitivity of different applications which may pose hard or soft delay deadlines. Especially the problem of energy efficient transmission subject to delay constraints over wireless networks has been studied extensively over the past few years. Traditionally, the problem of data scheduling has been considered at the medium access control - MAC (i.e., packet) level which considers short time intervals of the order of few milliseconds at most. On the contrary, hereafter we focus on message transmission which is inherently very elastic and can be delayed for couple of seconds or even up to few minutes. Examples of such messages are e-mails, FTP (File Transfer Protocol) data transfers, updates of social networking portals, data exchanges over P2P networks and RSS feeds to mention just a few. As navigation systems is a standard equipment on most vehicles, it becomes feasible to track the mobility pattern of vehicles which can assist for data transmission scheduling over long time intervals. When a vehicle moves along a road segment, the factors that will influence the wireless message transmission are the channel conditions, distance to the serving BS, the speed of the vehicle, and the population of Primary Users (PU) that would frequently change from time to time. Therefore, the decision to be made for a SU (vehicle or pedestrian in this case) is how to factor all these issues (sometimes conflicting) and choose an optimal duration and time to transmit a certain amount of data (messages as we interchangeably use hereafter) to the BS.

1.1 Energy Cost on Mobile Applications

The emerging demands of the mobile world are growing very fast with data traffic increasing significantly over the last few years. To quantify that, Cisco recently claimed that global mobile data traffic has a 69% growth in 2014, which reached 2.5 exabytes per month at the end of 2014 [2]. With further proliferation of smartphones and tablets, the expectation is that users will utilize a significant number of Internet applications which require an always-on connectivity (such as emails, updates to social networks, etc.) and have significant delay elasticity. In this situation, it is critical to develop innovative strategies to manage the wireless embedded systems efficiently, thereby prolonging the battery lifetime and enhancing overall user experience. In [3], the authors provide estimations of online power consumption of energy-constrained portable embedded systems (smartphones) to help end users achieve energy-efficient operation of the smartphone. The work in [4] explores a technique to achieve an increase of the battery lifetime for smartphones by shutting down its wireless network card when it is not being used. Moreover, the research in [5] examines in a detailed manner the top six energy consumption components on a smartphone and reveals that data downloading via cellular network and WLAN are two most important energy consumption functionalities. The power consumption of a WiFi radio accounts for a significant proportion of the overall system power in mobile devices. For a connected mobile device in idle mode, which the LCD and backlight are consuming zero power as they are turned off, the wireless interfaces consume approximately 70% of the total power [6] [7].

1.2 Device to Device Communications (D2D)

Device to Device (D2D) Communications is a feature which has been introduced in the 3GPP Release 12. D2D communication has been considered as an underlay to an LTE-A cellular network. Devices are allowed to be engaged in direct communication with the network having the control in terms of interference management and resources used [8]. In this context, both the cellular network and the D2D communication use the same LTE resources. Therefore, D2D communications can be considered as an enabler for the delaying message transmission which can be used to relay information to another device or to the base station.

1.3 5G Wireless Communications

Currently mobile operators trying to cope with the high demand of Internet applications in cellular networks that utilize carrier frequencies that range between 700 MHz and 2.6 GHz. Available spectrum at these carrier frequencies can be deemed as rather limited and as a result in order to increase aggregate transmissions rates to cope with the ever increasing demand there is a need to move higher in the spectrum. To this end, wireless technologies for 5G, or Beyond 4G as it is also called, envision the use of the current very much underutilized milimeter wave (mm-wave) frequency spectrum such as for example the use of 28GHz, 38Ghz [9] or even the unlicensed 60GHz as envisioned by the Wireless Gigabit Alliance [10] [11]. Clearly, at these frequencies signal attenuation is significant and as a result high speed broadband access to the Internet can be considered only for pico cells with radius of up to 200 meters [12]. In these scenarios delaying transmission of elastic user traffic based on the proposed set of solutions can allow for better utilization of the very high speed mm-wave links since it can allow users to refrain transmission until they are within the coverage area. Another important benefit stemming from the use of delaying message transmission until the user is closer to the access point is that in mm-wave spectrum due to the significant path losses the energy gains that can be achieved by delaying the transmission are even greater compared to current operating frequencies in cellular networks. By inspecting the Friis Law for free space path loss it can be seen that when moving from 3GHz to 30GHz path loss increase by 20 dB.

Also in these carrier frequencies the Power Amplifier has an efficiency of less than 10%; therefore energy consumption is a key issue. Consequently, by utilizing the elasticity of Internet application and the inherent mobility of users delaying message transmission is well fitted to be utilized in envisioned 5G wireless networks that are based on the use of mm-wave frequency spectrum.

1.4 Motivation

1.4.1 Availability of TVWS

Considering that an increased number of countries world-wide are permitting operation of cognitive radio systems in the vacant terrestrial analog TV transmission - TV White Space (TVWS), this would enable new possibilities for provisioning further capacity increase for wireless broadband and multimedia services [13]. In the US, the unoccupied TVWS has already been filled up with unlicensed users without significant interference to TV viewers, while Ofcom is determined to permit TVWS for unlicensed use by checking with a database in the UK [14]. Under this framework, when a number of SUs (Secondary Users) are using cellular networks within the coverage of a TVWS master, it is possible that the SUs would prefer the TVWS connection over the cellular networks in terms of cost and RF coverage. The SU connection can utilize the means of spectrum sensing or alternatively contacting a trusted geospatial database that records the information regarding PUs (Primary Users) occupation with a specific location and time duration, prior to message transmission, to determine available spectrum at a given location [15]. In this scenario, predicting the future location and the path of mobility of SUs (Secondary Users) is another challenging issue in White-Fi networks.

Spectrum is a fundamental resource in wireless networks since achievable data rates are linearly dependent on the available spectrum in use. It is widely acknowledged that the utilization of the spectrum can be deemed sparse and shows significant spatio-temporal variations. Exploiting these spectrum white spaces, as they are called, has been the focus of Cognitive Radios (CRs). CR networks encompass essentially two core functionalities, namely, spectrum sensing and spectrum management [16]. Spectrum sensing functionalities are responsible to detect unused spectrum and a plethora of different methods have been considered to do so which can be broadly taxonomized in cooperative and non-cooperative techniques reflecting whether the detection is taking place by the node itself or by utilizing other nodes in the network. Spectrum management on the other hand can be considered as the decision engine (mode of operation and parameters that will be used for transmission) on how to optimize the use of spectrum based on the information available from the spectrum sensing functionalities.

1.4.2 Spectrum Management

In essence, CR networks pose the potential of improving spectrum efficiency by allowing the low-priority SUs to temporarily seek the wireless spectrum that is licensed to different organizations (PUs) [17]. As soon as the PUs emerge in the frequency channels, the SUs must vacate the licensed bands. Hence the SU connections would be interrupted by the stochastic nature of the PU traffic. Consequently, the SUs should firstly estimate the channel availability by probability analysis based on PUs' historical traffic information or spectrum sensing. We assume that the SUs can achieve a perfect channel estimation about the PU connections from the Base Station (BS) or other wireless providers. In accordance with the information of channel utilization, SU connections can decide which frequency channels would be available without deteriorating the quality of the PU connections.

Queuing theory is a natural mathematical tool to analyze system performance by considering stochastic arrivals and departures as is the case in wireless networks. As such, queuing theory has been applied in significant volume of research efforts within CR literature. In [18], by considering a two-dimensional M/M/N/N queue, the authors have derived the mean number of the PUs and SUs with their blocking probabilities respectively. However, they focused merely on the situation where the SU connections would be immediately dropped if all frequency channels are occupied by PUs in Secondary Users Cleared (SUC) mode, and the situation that the PU and SU connections have equal priority in Secondary User Equality (SUE) mode. In our model, we assume that there are M/M/K/L queue systems for the SU connection, in which the SU message could be buffered.

1.4.3 Multiple Interfaces on Smartphones

A portable-size tablet terminal that enables radio communications in TV bands (470-710 MHz) has been recently developed based on the existing WLAN systems (IEEE802.11b/g) by inquiring a TVWS database. The device which has been developed by NICT can automatically select the optimal frequency for connectivity between the TVWS and the nominal 2.4 GHz bands by avoiding potential interference to other TVWS transmissions based on the information available from a TVWS database [1]. As modern smartphones and tablets envisioned to have multiple wireless interfaces enabled (Wi-Fi, 3G, LTE and TVWS interface), the energy cost for data transmssion will vary significantly due to different characteristics of wireless interfaces and distances with the corresponding access routers. The energy consumption for wireless transmission has been analyzed for delay-tolerant applications over 3G, GSM and Wi-Fi(802.11b) in [19]. A basic CoolSpots model is proposed to enable a seamless switching among radio interfaces of mobile devices such as WiFi and Bluetooth, in order to decrease overall power consumption while maintaining enough bandwidth to support active mobile applications [6] [20]. Since multi-mode network cards are becoming increasingly affordable, a wireless access router device is implemented in [21], which can be placed in moving vehicles to enable high-rate data access service. In this case, multiple wireless access links are utilized to aggregate wireless bandwidth and provide local users with a more reliable access network than the service provided by a single cellular networks link. Hence significant benefits can be achieved within the different coverages offered by several mobile operators and technology access. It can be in addition assumed that an infrastructure of roadside units or access points (AP) have become ubiquitously available in urban areas [22] [23] [24] [25]. Many telecom operators and local governments around the world are planning to offer city-wide deployments of Wi-Fi access as part of the basic infrastructure [26]. In [22], deploying infrastructure APs can effectively assist the data delivery for inter-vehicle communication. A novel routing protocol based on existing Wi-Fi APs and navigation system has been presented for large size file transmission in vehicle-to-vehicle transmission [23]. A system is proposed [27] to let a mobile device track connectivity quality as its user moves along similar paths every day with fixed habits. Without requiring GPS hardware and extensive centralized infrastructure, the proposed system can effectively

[1] www.nict.go.jp/en/press/2013/08/28-1.html

forecast wireless connectivity for mobile devices by maintaining a personalized mobility history on mobile device and tracking the APs encountered at different locations. The experimental results upon real-world usage indicate that even with one week training period, these forecasts are sufficiently accurate to improve performance of mobile devices and significantly reduce power consumption for several mobile applications. In addition, for the cases where the mobility of vehicle and pedestrian can be predicted, the roadside APs can evidently improve the average wireless throughput for file delivery by estimating the signal strength of APs along the familiar route from historical RF fingerprint data statistics [25]. The data collected in [28] showed that, during half year period, 99% and 49% of daily life on average for participants was under cellular networks and accessible Wi-Fi network respectively within urban area. Based upon that, a combination of short and long term Wi-Fi condition is investigated in [28], which provides a theoretical analysis on data transfer via multiple wireless interfaces by leveraging the complementary energy profiles of Wi-Fi and cellular network. As 33% of total mobile data traffic was offloaded onto the fixed network through Wi-Fi or femtocell in 2012 [2], to predict the location/trajectory of SUs and take advantage of high-speed hotspots (Wi-Fi, TVWS) outline another set of challenging issues in this scenario.

Meanwhile, a significant proportion of mobile applications are inherently delay-tolerant, such as social networking, email client, file downloading and firmware/security updates. Hence, it is possible to artificially delay data transmission until a high-rate and/or more energy efficient Wi-Fi or TVWS hotspot become available. The proposed scheme follows previous line of research in the area of energy-delay trade-off algorithms for delay-tolerant Internet applications [29]. The focus in this case is to utilize a wide spectrum of available wireless connectivity especially focusing in the area of small cells utilizing both Wi-Fi and TVWS networks.

1.4.4 Overall Cost on Embedded Systems

Battery capacity is not improving with the same pace as data rates, CPU cycles, available memory devices and display technologies on the terminal side. In this situation, it is critical to develop innovative strategies to manage the wireless embedded systems efficiently, thereby prolonging the battery lifetime and enhancing overall user experience.

In this research, different applications in portable embedded systems will be characterized to provide a practical model so as to reduce the overall energy cost of applications over wireless networks.

1.5 Contributions

The contributions of this thesis are briefly stated below.

1. Under the inherent stochasticity of available of transmission opportunities, the challenge is to select an optimized time duration to launch the data transmission in order to minimize the overall energy cost while satisfying the information delay constraint. To this end, A novel relaying scheme to deal with the problem of energy-delay trade-off, based on optimal stopping programming (OSP) in section 5.3, is proposed, whilst exploring different candidate solutions for wireless transmission. The problem of delay-tolerant data transmission scheduling over Cognitive Radio (CR) enabled wireless networks is considered with special emphasis on high mobility users (vehicles) in Chapter 5.

2. The methods for load balancing of SU traffic are studied in subsection 3.1.3, in accordance with the information of channel utilization, mobile systems can decide which frequency channels to utilize without deteriorating the quality of service of experience of the Primary connections. An M/M/K/L queuing system is modeled to analyze the performance of the SUs competition under various traffic blocking thresholds and queuing delays. By optimizing the frequency channel utilization the number of SU connection that can be accommodated simultaneously is derived. The model based on M/M/K/L queue can delay message transmission and make the use of available white spaces to reduce the energy consumption without affecting PU traffic.

3. The model is proposed to evaluate and manage the associate energy cost of streaming multimedia content over wireless networks in digital devices by incorporating and utilizing the inherent mobility of the users as presented in section 5.4. Given the fact that popular video clips are likely to be watched multiple times by a mobile user [30], a methodology is introduced to seek optimal solutions for the wireless

Internet access scheduling by taking the effect of stochastic locations of of base stations (BS), Wi-Fi APs, and energy consumption of storage devices. After an estimation of the energy cost in different locations and corresponding probabilities of customer demands, a set of policies are introduced which effectively avoid areas with relatively higher energy cost for the elastic streaming media content, thus prolonging the lifetime of digital devices.

4. Once smartphones are equipped with multiple wireless air-interfaces, and mobile applications are willing to tolerate some delay in exchange for high-quality transmission in order to extend smartphone lifespan, the proposed schemes could delay some file delivery and effectively avoid the areas with higher transmission cost so as to improve battery longevity. The optimal wireless Internet access scheduling schemes also take into account spectrum availability in subsection 3.1.3 and the effect of stochastic access duration of wireless hotspots in subsection 4.2.2.

1.6 Thesis Structure

The main aim of this research is to design a practical model to evaluate and manage the associate energy cost of Internet applications over wireless networks in digital devices by utilizing the inherent mobility of the users. In this framework, the optimal wireless Internet access scheduling model is required to reduce the overall energy cost that related to user locations, mobility pattern, and applications delay tolerance, which also takes into account spectrum availability and the effect of stochastic access duration of wireless hotspots, while satisfying the delay constraints of mobile applications. Therefore, by optimizing the frequency channel utilization for simultaneous SU connections, a novel relaying scheme is proposed, based on delay elasticity of mobile applications, in order to reduce energy consumption by utilizing spatio-temporal characteristics of users.

The structure of this thesis is specified in Figure 1.1. An overview of the existing research, such as Delay Tolerant Applications in the literature and spectrum availability discussion, is given in Chapter 2. The framework and components of multiple wireless accesses are introduced, and the distribution of SU traffic loads and an M/M/K/L queuing system model are analyzed in Chapter 3. To instantiate user behaviors under the architecture of multiple wireless interfaces in Chapter 3, the stochastic characteristics of user mobility are described in Chapter 4, which include direction changes and variable speed of the mobile users. Finally, the characters of mobile applications and detailed energy cost in mobile devices are introduced in Chapter 5 to reduce overall energy cost base on mathematical model building. The conclusions and directions for future work are presented in the final Chapter 6. The details of each chapter are explained as follows.

- Once mobile devices are equipped with multiple air-interfaces allowing them to connect to LTE, Wi-Fi and White-Fi, they could switch among these networks to energy efficiency strategies to achieve significant energy gains. Chapter 3 describes the framework with multiple wireless accesses (hotspots). Especially, as a Secondary User, if the mobile applications should effectively seek and use any licensed spectrum bands as long as they do not cause interference to the PUs, they could significant lower itself energy cost for wireless transmission. Hence, in Section 3.1.3, an M/M/K/L queuing system is modelled to analyze SUs competition under various traffic blocking thresholds and queuing delays and provide load-balancing by utilizing the distribution of SU traffic loads.

- When the users mobility can be well predicted, systems could have a prior knowledge of the mobile user's destination. In realistic wireless networks, mobile users may choose to take a longer route to the destination. Consequently, the mobility modeling should include changes in the direction and speed of the mobile users. Therefore, the stochastic characteristics of user mobility are introduced in Chapter 4 which should include direction changes and variable speed of the mobile users. Vertical handover enables mobile devices switch between high-coverage access and high-rate access for cost-efficient wireless transmission. Hence, a framework of heterogeneous networks is introduced to provide the seamless wireless connections for mobile users in subsection 4.2.1. Cell residence time in different wireless hotspots is analyzed to search the optimal schemes that can provide better performance in terms of mobile application requirements, wireless transmission and switching cost of radios in embedded systems in subsection 4.2.2. In addition, as the power consumption on wireless radios accounts for a significant proportion of overall energy cost in mobile systems, a novel strategy is proposed in section 4.3 for battery lifetime extension based on delay tolerance of mobile applications by intellectual use of the available high-rate roadside wireless accesses.

- Mobile platforms accounted for 60% of total digital media time spent. And mobile applications made 51% of all digital media time spent in May 2014 in the United States [31]. Hence the strategies, which manage the wireless embedded systems efficiently, are critical for mobile battery lifetime extending. The analysis for energy saving in mobile device in Chapter 5 incorporates energy cost on embedded peripherals such as CPU, Graphics, backlight, storage devices, and delay cost for delay tolerant applications. In section 5.3, a novel scheme based on optimal stopping programming is proposed to deal with the trade-off of energy-delay. Through a wide set of numerical investigations we shed light on the possible achievable gains. Apart from that, recent research indicate that mobile users may want to access the same popular video clips several times which can be energy inefficient if it is always streamed to the users the policy regarding download and store trade-off, by taking into account the probability of re-using the content, the energy cost on storage devices and the energy cost on wireless streaming, is introduced in section 5.4. Finally, in section 5.5, by characterizing different mobile applications, a practical model is proposed to reduce the overall energy cost of mobile terminals.

- Chapter 6 summarises the thesis contributions and presents final conclusions. Further it is briefly explained possible avenues of future directions for energy-efficient wireless transmission within the DTNs and cognitive networks in section 6.3.

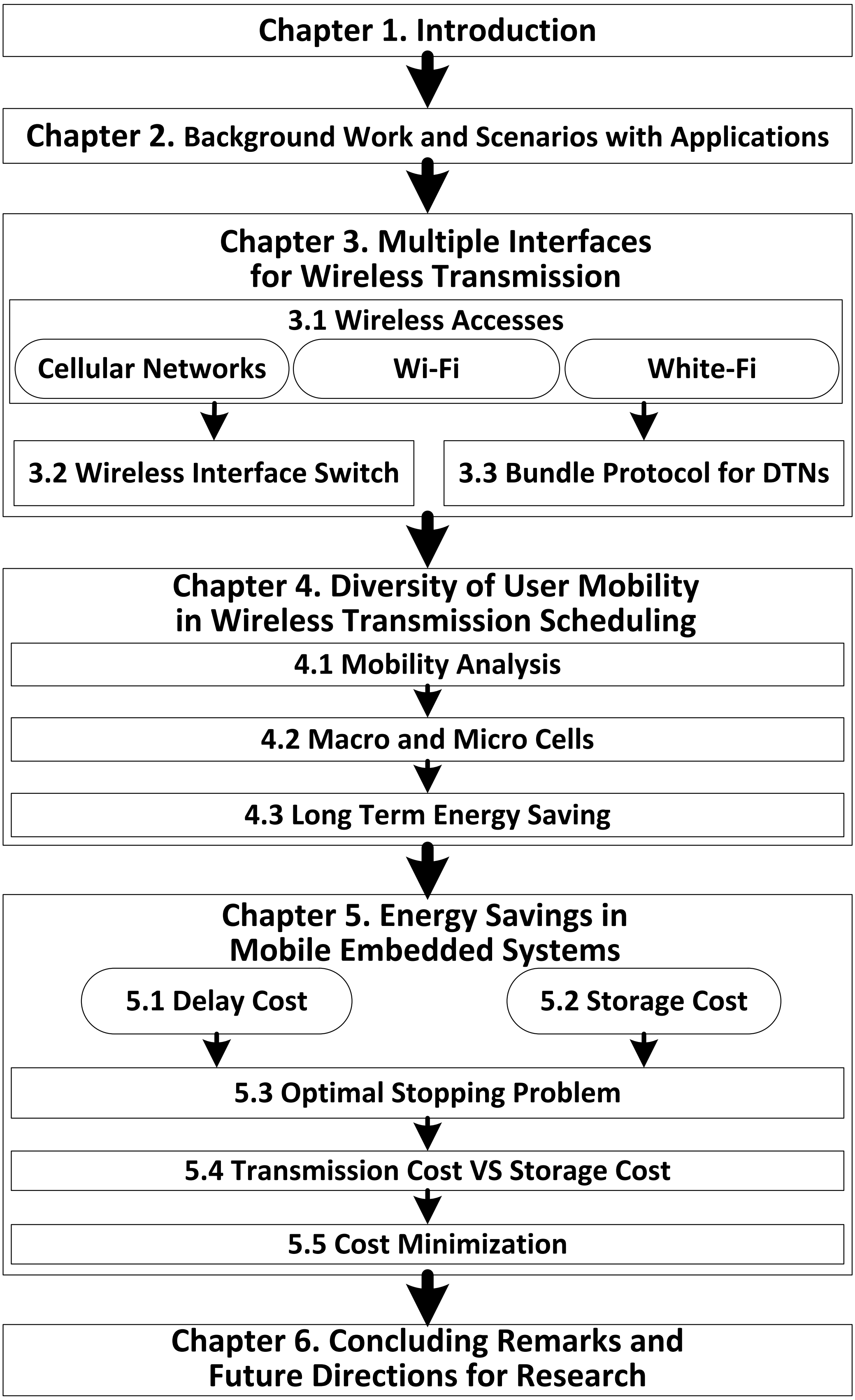


FIGURE 1.1: Thesis Structure

Chapter 2

Background Work and Scenarios with Applications

In Chapter 2, a brief survey of ideas and technologies from other researchers will be compiled, which are tightly related to the problem of energy-efficient delay-tolerant networking. A comprehensive review of two important technologies for emerging and future cellular networks, namely Cognitive Radios (CR) and Delay Tolerant Networking (DT-N) are introduced in Chapter 2, which have been considered in isolation until now. An overview of how to capitalize the delay tolerance of various applications are introduced, followed by the introduction of the distribution of PU and SU traffic loads in cognitive networks. Finally, this chapter also outlines previous related research in the area of vehicular communications, opportunistic networking, and energy saving techniques.

Hereafter the focus is on Internet applications which can tolerate significant delays without deteriorating the experience of the end user. To this end, the consideration is on highly delay-tolerant traffic which can tolerate delays from few seconds up to few minutes. Traffic with these characteristics are e-mails, updates of social networking portals, message/file exchanged via the FTP, RSS (Rich Site Summary) feeds, non-real time video streaming and OS (Operating System)/firmware updates to mention just a few. We can utilize their inherent characteristics to significantly save the energy cost for embedded systems of mobile nodes. Note at this point, and as have been mentioned previously, the increased usage of smartphones and the rich ecosystem of Internet applications are having a severe effect on the recharging cycles of devices due to the increased levels of energy consumption and limitations of battery technology. As the infrastructure of hotspots for Wi-Fi (IEEE 802.11) and White-Fi (IEEE 802.11af) interfaces are becoming ubiquitously available in urban areas, and mobile devices are equipped with multiple air interfaces, they could switch among these networks for better performance and lower energy cost. In this case, the energy usage in modern devices for transmitting a fixed amount of data could differ drastically due to the significant difference on the achievable data rates on these radios. In addition, channel conditions change according to user mobility and spatial characteristics of the channel. Therefore, predicting the future location and the mobility path of SUs (Secondary Users) is another challenging issue in White-Fi networks. Furthermore, apart from the availability of the primary channels, the mobile nodes have to compete with other SUs to seek an optimal time duration for wireless transmission. Consequently, a virtual queuing model based on an M/M/K/L system is designed to analyze the optimal population of SUs to be served in the system to minimize the energy consumption of message transmission for the delay-tolerant applications. Wireless nodes (SUs) gather PU traffic information from a historical database in order to predict over a short-term the traffic pattern of PUs. In the database server, there are two types of information about primary channels. One is the 24-hours traffic characteristics of different channels across several months [32], and the other is the noise power level in different channels that updated by SU devices constantly. A SU has to send a query to the database server for free channels to transmit. According to the available channel information at the same time slots in previous days and long-term statistics regarding channel availabilities, the database server will certificate the noise power levels by comparing to a threshold. Then the best candidate channels for the inquired SU will be determined.

2.1 Delay Tolerant Networks

Significant volume of research efforts have been placed on energy efficient data transmission for delay tolerant applications, especially the trade-off between transmission cost and time delay over wireless networks [33] [34] [35] [36] [37] [38]. The work in [33] deals with the problem of packet scheduling with deadlines within a pre-defined time window of length T. Based on that, the authors in [34] explore the energy-efficient packet transmission with individual packet delay constraints. They indicate that when packet inter-arrival times are independent and identically distributed (i.i.d.), the optimal transmission durations of two symmetric packets are also identically distributed. According to this property, a flexible energy and delay tradeoff is illustrated under various individual packet delay constraints and bandwidth efficiencies.

In [35], the authors consider a delay constraint for each packet and reveal the relationship between reliable transmission rate and QoS requirements, while a dynamic programming based algorithm is introduced to acquire throughput maximization and energy minimization according to different channel qualities of a fading channel with time constraints [36]. The work in [37] further investigates the problem of energy-delay trade-offs under dynamic traffic loads and user populations. The target-set selection problem has been studied in the emerging Mobile Social Networks for traffic offloading by delaying the delivery [39]. In [40], a framework is proposed to investigate the trade-off between the amount of offloaded traffic and the users' delay tolerance over a 3G network.

For delay sensitive traffic there has been an enormous previous research both applied and theoretical within the general area of optimal job scheduling with deadlines [41]. There has also been significant volume of previous research work on energy efficient scheduling for wireless packet transmission.

Today advanced mobile devices (smartphones and Internet tablets) are commonly used both at public or private places, which are used for email, accessing RSS news feed, watching streaming video, listening music/radio - to just name a few the prominent features. This trend has caused a considerable increase in mobile Internet traffic. In terms of wireless access connectivity, while being mobile users are more likely to access the Internet via cellular macro-cells connectivity due to the ubiquitous coverage, which is in turn creating immense pressure on the limited bandwidth of cellular networks. On the other hand, if the information regarding the location and coverage of high-rate hotspots (small cells) can be easily obtained, Internet access via these hotspots could be another option for users in response to the traffic pressure on cellular networks. However, due to the noncontinuous coverage of small cells, there has to be a delay for mobile users to move into the coverage area of such high-speed hotspots (Wi-Fi, White-Fi). In this case, the energy consumed for highly elastic Internet traffic could be significantly reduced, while respecting user-specified and applications delay-tolerance.

In [42], a joint study of 3G and Wi-Fi characteristics is conducted. It is introduced that relatively costlier cellular networks provide ubiquitous connectivity, while free Wi-Fi service is intermittently available. Therefore, the aim of this research is to offload cellular data transmission as many bytes as possible to Wi-Fi while meeting application-specific requirements.

Table 2.1: Features of Delay Tolerant Applications

Type	**Percentage of Data Consumption**
Mobile Web/Data/VoIP	36% ∼ 19% [2]
Mobile Video (Youtube)	55% ∼ 72% [2]
Mobile Audio (Podcast)	8% ∼ 7% [2]
Mobile File sharing (dropbox)	1% ∼ 2% [2]

According to the different features and user requirements, the mobile applications can be categorized into several types. Table 2.1 compares the existing Delay Tolerant Applications in current digital devices. This table also show applications that the proposed schemes as detailed in later sections can utilize their delay elasticity to reduce energy cost. First of all, Table 2.1 presents the percentage of the top mobile applications, in which the most popular mobile application is stored video streaming that accounted for up to 55% of total mobile data traffic by the end of 2014. And Cisco predicts that 72% of the global mobile data traffic will be video streaming by 2019 [2]. The study in [43] defines the threshold of delay value experienced before the actual video playback as 1, 5 and 10 seconds. So we set a minor delay (10 seconds) for the stored video streaming applications based also on user requirements. As the mobile advertisements have already been embedded in current mobile platforms, during the delay gap, the advertisements that downloaded locally can be played to fulfil the user experience. Secondly, this is followed by mobile Web/Data/VoIP. Some applications, such as e-mail, social networking and RSS news feeds periodically exchange a portion of data with a remote server via wireless access. For instance, podcast content on mobile devices will be synchronized many times a day, thus we could utilize its delay-tolerance characteristics to reduce the download cost [44] [45]. In [19], the maximum delay of e-mail has been set to 10 and 15 minutes in order to reduce the energy consumption of download, while the delay threshold of news feeds has been selected as 10 minutes for the energy reduction purposes. Recent estimates also confirm that the rate of music streaming and email use will increase along with the the growing popularity of smartphones [46]. In [47], a system architecture for delay-tolerant wireless networks has been proposed and focuses on the public content distribution such as software updates and news distribution.

More importantly, Intel has designed a DTN architecture for mobile environments that lacks continuous connectivity [48]. In their proposed architecture, a Bundle gateway located above the TCP/IP protocol stack can provide a store-and-forward data mechanism. Based on the store-and-forward approach, DTN capable phones have been implemented as a carrier for DTN bundles on the Android platform [49]. The research hereafter is motivated by the fact that a significant proportion of mobile applications are inherently delay-tolerant, such as updates in social networking, emails, RSS feeds and weather updating to mention just a few. Hence it is possible to delay data transmission until high-rate Wi-Fi or White-Fi hotspots become available. It has come to our attention that many previous studies focus on the energy-delay tradeoff algorithm for delay-tolerant smartphone applications [29]. Consequently, our approaches focus on the energy-delay tradeoff research. For the delay-tolerant multimedia content, we recommend mobile systems delay some file delivery and effectively avoid the areas with higher transmission cost so as to improve battery longevity.

2.1.1 Video Streaming

In [50] the authors analyzed the popularity distribution and the statistical properties of requests for user-generated content (UGC) online video. According to the Pareto Principle, about 10% of the top popular videos account for nearly 80% of the views [50], which means that caching a small set of multimedia video could make an appealing performance enhancement for the content with high download frequency. They further show that caching the most popular content can offload 50% server traffic. In [51], the YouTube traffic in campus network has been monitored for long-term period, and it can be observed that over 50% of the video requests and corresponding bytes transmission were for previous request, which reveals that caching in local storages has the potential to reduce bandwidth demands for video content.

2.2 Availability in Cognitive Radio Networks

A common technique for channel estimation has been designed to analyze the PU traffic characteristic from available long-term observations/statistics [32]. In [52], authors design optimal sensing strategies via a model assuming that the PU transmissions are unslotted as a continuous-time Markov chain while the SUs are slotted to sense the frequency channels. On this basis, Noh et al. propose a stochastic multichannel sensing scheme based on traffic information and sensing history [53].

In the meantime, a preemptive priority queuing system has been utilized to analyze the mean system dwelling time of the SU traffic and the blocking probability for real-time SU connections [54]. In [55], the authors analyze the queue lengths and average queuing delay of the SUs based on Poisson distribution of the SUs. In [56], a Dynamic Strategy Learning (DSL) algorithm relied on the priority queuing systems including the SUs and the PUs is proposed for the delay-sensitive multimedia applications in order to maximize the user's utility function.

There have been several prior works on dynamic spectrum access, especially channel selection schemes. In [57], a dynamic protocol is introduced for dynamic spectrum allocation (DSA) with load balancing for SUs. They assume that the PU traffic detection and SU connections blocking in these channels from the allocation game has been solved already, therefore they just focus on the load balancing of SU traffic. Their algorithms are modeled as a so-called balls and bins problem of congestion game, where balls representing the CRs are assigned to several bins representing physical channels of the radio spectrum and SUs may reassign its load in a round based fashion dynamically. The system can rapidly converge towards an approximately balanced state in which all CRs will sustain cost below a certain threshold parameter. In [58], CR channels are logically divided into multiple bands in order to achieve excellent load balancing performance, where each band consists of a group of channels.

In [59], in order to achieve better channel utilization, SUs need to make a decision at each time slot to access if the current achievable throughput can be improved by utilizing another physical channel based on channel sensing and estimation. This channel exploitation problem can be modeled as a finite-horizon optimal stopping problem between the expected increasing data rate of SU and channel sensing cost. Opportunistic scheduling policies are developed in [60] to maximize the throughput utility of SUs with general interference and mobility models under maximum collision constraints with PU traffic, by designing collision queues that ensure reliability constraints. As the available spectrum could be changed when mobile users are moving from one place to another, continuous allocation of spectrum is another major challenge for mobile users [61].

My research analytically derives a scheme that allows the SUs to decide the transmission duration in order to lower the energy consumption. Once the mobile devices equipped a dual or multi mode antenna that is connected to cellular networks and other networks like Wi-Fi and White-Fi, the SUs could switch between these networks to seek and use any licensed spectrum bands as long as they do not cause interference to the PUs. Since the SUs can require a trusted database for the information traffic of PU connections, we could predict the PU traffic within a specialized relatively long-term duration. Meanwhile, a virtual queuing model based on M/M/K/L system is designed to analyze optimal SU number to be served in the system to minimize the energy consumption of wireless transmission.

2.2.1 Queuing System

Many different facets of queuing theory have been utilized in the CR literature to analyze the different priorities schemes for PU and SU traffic. The research in [62] models the data transmission and channel sensing problem as a queueing system in order to make opportunistic use of spectrum channels which have been licensed to some primary network. In [63], a slotted transmission system and an infinite queue is assumed for both PU and SU traffic. SUs sense the channels in each slot, and transmit a packet from the corresponding queue if detecting an idle time slot. In this context, SU link has been considered as a "transparent" relay for the PU traffic, if it does not affect the stability of the PU queue. In [64], the preemptive resume priority (PRP) M/G/1 queuing network model is developed to characterize the channel usage of CR networks. Based on this queuing model, they characterize the spectrum utilization behaviors of each channel with different arrival rates and service time distributions of PUs and SUs. The research investigates the effects of spectrum hand-off delay on the extended data delivery time of the SU connections. Based on the proposed PRP M/G/1 queueing network model, a suboptimal greedy target channel selection scheme is proposed in [65] to select optimal target channels. SUs must execute a sensing procedure to check availability of the current operating channel before data transmission in a slotted-based CR network. According to spectrum usage interactions between PUs and SUs, the SUs can resume the unfinished transmission by finding suitable idle channels with with shortest delay of spectrum handoffs. In each time slot the cognitive user, SU can can only sense one of N parallel primary channels to exploit transmission opportunities for data accumulated in its queue in [66], which primary transmissions are statistically independent and are modeled as discrete-time Markovian on-off processes.

A significant volume of research took the approach of categorizing users into K priority classes. In that case, the highest priority class C_1 is always reserved for the PUs in each channel and SUs occupy the rest of $K-1$ priority classes $(C_2, C_3, \cdots, C_k)$ [56] [67]. In [56], an M/G/1 model is adopted and each SU maintains several physical (or logical) queues for the various frequency channels. The channel selection decisions are based on priority queuing analysis that considers the dynamic channel conditions, traffic characteristics, and the competitors' behavior. Hence, the SUs could efficiently adapt their channel selection strategies for the delay-sensitive multimedia applications in order to maximize the user's utility function. In [67], they also model the traffic as an M/G/1 preemptive priority queue with channel conditions and one PU and multiple SUs competing for the same frequency channels. Vacating the channels for the PU traffic, SUs would sense the remained channels and time slots (spectrum holes) according to a simple First-Come-First-Serve (FCFS) rule. Moreover, they place the emphasis on the delay and throughput of the SU traffic instead of stability.

An M/M/1 time-varying queuing model is introduced to generate an accurate temporal and frequency characterization of CR networks in [68]. In this case, the single server is assumed to be a centralized BS and spectrum occupancy and availability for SU connections is based on a statistical model. In [69], the authors design a model that PUs occupy many licensed channels with one channel allocated for SU traffic. The SU traffic subsystem is modeled as an M/M/1 queue with a processor sharing (PS) policy. Meanwhile, the SUs would sense the PU licensed channels in order to find a free channel for transmission. As a result, the traffic of PU license channels could be modeled using an M/M/K/K queue. PU traffic in three different levels of priority compared to

SU connections (perfect, partial and no priority) has been analyzed in [18], where an M/M/N/N queuing system is utilized to analyze PU and SU traffic of arrival process and service rates.

In [54], the preemptive priority queuing model is designed to analyze the SU traffic of CR system. The distributed non-real-time SU traffic enters a queue buffer in First-In First-Out (FIFO) manner when the server is busy with PU or earlier SU connections. In non-real-time situation, SU traffic are allowed to be buffered in queues. Meanwhile, if there is no server available, the real-time SU traffic are immediately discarded. They show the simulation results regarding system dwelling time for non-real-time SU connections, blocking probability and forced termination probability for real-time SU traffic under different SU arrival rates. In [55], authors analyze the queue lengths and average queuing delay of the SUs based on Poisson distribution of the SUs.

2.2.2 Spectrum Availability

As the vacant TVWS spectrum are permitted to be used in several countries, and it is without doubt that this new available spectrum will unfold new possibilities in data transmission with strong potential to further decrease overall energy consumption. From the research in [70], the percentage of locations with non-zero number of available frequency channels is 64.7% under the ECC rules. In other words, for a SU, there is a high probability that at least one frequency channel in TV white space could be available for wireless transmission. Therefore, in this scenario, what is required is an estimation of the probability that the SU device is within the coverage of a Wi-Fi like AP. We further analyze the number of Wi-Fi like APs and the location of these APs in a specified urban area. Opportunistic access to the licensed spectrum will be interruptible in the sense that cognitive users have to cease transmission immediately and relocate to a new band as soon as the primary user appears. In [1], the authors propose an effective spectrum decision scheme, which can evenly distribute the traffic loads of SU connections to multiple channels, thereby reducing the average overall system time compared to the non-load-balancing scheme. Furthermore, the paper discusses two kinds of spectrum decision schemes: one is sensing-based with the objective to determine the optimal sensed number of candidate channels for channel selection, the other is probability-based that takes the traffic statistics of both PUs and SUs. Different schemes for TVWS access can be taxonomized as follows.

1) In the database server, there are two types of information about primary channels. Firstly, the 24-hours traffic characteristics of different channels in several months [32]; Secondly, the noise power level in different channels that is updated by the SU devices constantly. When a SU needs a channel to transmit, it sends a query for available channels to the database server. Based on the historical information about unoccupied channels at the same time slots in previous days and long-term statistics about channel availabilities, the database server will certificate the noise power levels by comparing to a pre-defined threshold. Then the server will determine the best candidate channels for the inquired SU and inform about the channels with high probability to be unoccupied. In this scenario, we assume the SU has to periodically send the queries to database channel to ensure the best channel conditions (SU throughput) and vacate as the emerging of PU connections. Intuitively speaking, reducing the time interval that the SU sends one query to the server, it will guarantee the SU to perform successful transmissions and avoid interference for PU traffic. Nevertheless, it will increase the transmission cost for the SU. As a result, this channel exploitation problem can be modeled as a trade-off between the query interval and energy cost of SU. In [71], the authors design a minimum collision rate and minimum handoff rate algorithm to maximize spectrum holes utilization of the PU channels for optimal SU throughput on the basis of satisfying constraints of collision tolerable level of PU channels. These two channel selection schemes are based on spectrum holes prediction from past observations. In [72], a opportunistic channel selection scheme is proposed with the aid of statistical traffic pattern of PU channels and traffic prediction techniques. According to the long-term statistical probability of each channel appearing idle in the next time slot, SU will firstly sense the channel with the highest probability of being idle. Likewise, a channel selection method has been devised in [73] to analyze the most probable unoccupied channels for SU traffic. When a database receives a SU query for an unoccupied channel, it will deliver stored historical information about channel availability. The database searches the channels that have been unoccupied at previous days from that time slot to some time in the future, while it ensures noise levels of the most probably vacant channels is lower than a threshold. As an extension of this research, in [74], all channels will be sensed and spectrum availability

of channels together with geolocation information will be updated in the channel history database by the most recent sensing results, which the information about spectrum use in different locations are stored in binary format. According to their prediction results, the proposed channel selection scheme can find out optimal channels offering the largest predicted idle time durations for SUs.

2) Regarding the sensing based scenario, the unlicensed users are allowed to transmit if they sense the licensed band to be unoccupied. In [62], the spectrum sensing approach adaptively schedules the sensing periods to improve the spectrum efficiency of CR channels. The principle followed is that the proposed schemes transmit data when the primary channels are in good conditions, while spectrum sensing is operated when the channels are poor. In this category there are two types of sensing techniques. Delay-sensitive applications over mobile devices prefer the proactive approach that cognitive devices will sense the spectrum periodically to maintain a list of available bands instantaneously, which would clearly incur increased sensing overhead. Meanwhile, if the applications are delay-tolerant, the cognitive devices favour the reactive schemes that sense the spectrum only when data are becoming available for transmission. Apparently, the on-demand sensing strategy is more energy-efficient. However, once the cognitive devices start to occupy one white space band, it has to proactively sense the TVWS spectrum at periodic intervals.

Under reactive spectrum access model, SUs will switch channels only after detecting PU traffic. Hence, there will be an unavoidable window of possible interference to PUs. Meanwhile, SUs would suffer from unexpected interruptions during transmission, making it extremely difficult to satisfy requirements of applications. With dynamic spectrum availability, SUs must monitor spectrum constantly and switch among channels to avoid disrupting PU traffic. Furthermore, since QoS and aggregated throughput of SUs could be dramatically degraded due to the burdensome cost in channel switching, it is more favorable to avoid the costly channel switchings [75].

(A) spectrum prediction switching

The process of channels can be modeled as a two-state Markov Chain birth-death process with death duration α_n and birth duration β_n. An ON (Busy) state represents the period used by PUs, while an OFF (Idle) state exhibits the unused period. The lengths of the

ON and OFF periods in each channel are independent identically distributed (i.i.d.) with mean value equal to α_n and β_n [72], while another popular model is that the durations are independently exponentially distributed [76] [77]. In the latter case, for the channel i, the mean duration of OFF period approaches $1/\lambda\alpha_i$. SUs can estimate future spectrum availability via spectrum sensing, by acquiring the knowledge of statistical property of primary spectrum's usage pattern. This knowledge is from the approach by observing primary channels over a long period of time or from assistance of a spectrum server. The sensing scheme in [74] is a periodic sampling process so as to determine the ON/OFF state of primary channels. The updated spectrum information about primary traffic in different channels will be stored into a channel history database for prediction of spectrum availability with OFF state. Instead of switching channels randomly upon unpredictable behaviors of PU traffic, a channel selection strategy is implemented in [75] based on statistically optimal channels in order to minimize the conflicts with PUs and maximize the probability of successful transmissions. Based on learning automata techniques, each SU dynamically maintains a probability vector of choosing each primary channel, and selects an available transmission channel accordingly. In [78], for low overhead and high likelihood of spectrum availability, SUs has to decide the optimal number of spectrum channels to be sensed if a decision is made in favor of seeking spectrum opportunities.

(B) random switching

In each time slot, the SUs have to scan all available channels or randomly select some of the channels if the number of channels is significantly large. In this case, the scanning and switching cost will account for a large proportion of energy consumption, thus increasing the overall transmission cost.

2.3 Heterogeneous Network

A heterogeneous-network deployment constitutes a promising concept of meeting future data-rate and capacity demands, which can satisfy high traffic demands by utilizing high-speed hotspots [79, 80]. In [81], Phantom Cells with high capacity are located within the coverage of LTE cellular networks to enhance the user throughput at Macrocell's edge regions. In the literature, a great deal of effort has been devoted to optimize the vertical handover among different radio accesses.

2.3.1 Interface Switch

The new generation of cellular technologies (4G or LTE) may have difficulties to satisfy the significant bandwidth demands of Internet access requirements of mobile applications on recent smartphones and tablets in the future. From mobile users' viewpoint, they would probably prefer the fastest possible data transmission over mobile device with long battery lifetime. In addition, mobile users would prefer cheaper networks such as free Wi-Fi, as typical cellular networks are billed in function of the number of transmitted bytes [82]. Therefore, cellular networks should utilize high-rate accesses to offload data for mobile subscribers [83] [84]. The key requirement for MNs is that they should be able to adapt its protocol stack to its user's requirements.

As the mobile operators start to have difficulties to carry growing wireless traffic over cellular networks, overloading to high-speed hotspots is being considered. Multipath TCP is proposed as an evolution of TCP in [82] which allows the simultaneous use of multiple interfaces while still preserving a standard TCP socket API to the application. Experimental investigations for Multipath TCP are presented to prove the feasibility of multiple radio accesses handover among wireless networks for real applications. With the ubiquitous deployment of cellular data networks, a system is designed to reduce the immense pressure of 3G networks by leveraging available WiFi connectivity without significantly hurting user experience of mobile application [42]. Cellular networks commit to provide always-on, ubiquitous data service connectivity with relatively low data rates, while Wi-Fi hotspots offer high data rates with limited coverage. The performance of wireless data services would be dramatically improved if users could seamlessly roam across these two networks. Therefore, two possible approaches are introduced for an integrated 3G/WLAN network architecture in [85], namely tightly-coupled and loosely-coupled interworking. By using qualitative analysis, it is showed that the latter is the preferred approach. Since compared to Wi-Fi, cellular networks require much lower power to stay connected but incur a more than an order of magnitude higher energy than that of Wi-Fi per MB transfer, an approach is proposed in [28] to achieve energy-efficient data transmission by selecting multiple network interfaces among increasingly available Wi-Fi networks and ubiquious cellular networks. Algorithms is designed to learn effectively from context information and estimate the probability distribution of Wi-Fi network conditions so as to decide whether or not to power up Wi-Fi interface for wireless transmission for the purpose of energy saving.

On the top of that, under such a mobility networking environment, the vertical handover among heterogeneous wireless networks has to be supported. During the process of vertical handover from one access to another access, the MN cannot transmit/receive data on its new point of attachment until the vertical handover ends. This can result in disruption of an ongoing mobile streaming and dissatisfaction of the mobile user, and higher user mobility leads to frequent handovers and service disruptions. Hence, it is important to evaluate the handover latency of mobility protocols [86]. A vertical handover Decision Algorithm for mobile phones is proposed in [87], which considers the power consumption of mobile phones using different wireless access technologies (UMTS and Wi-Fi) in a heterogeneous network environment.

In [88], an EIS (Enhanced Information Server) architecture is proposed to accelerate vertical handover procedures, which wireless channel conditions are estimated by exploiting spatial and temporal localities at the EIS to reduce the vertical handover latency among WiMAX, HSDPA and CDMA. The MNs periodically send the message including its location, RSS (Received Signal Strength) conditions, and the timing information to the EIS. Consequently, it is possible to estimate the channel conditions and determine the best PoA (point of attachment) for the MN, thus skipping time-consuming channel s-canning procedures for a handover trigger event. In [89], a low-complexity RSSI-based algorithm is proposed based on the implementation of a real testbed in IEEE 802.11 Wi-Fi and 3G networks, which the vertical handover operations are completely operat-ed by the mobile terminal. In this case, the handover latency is typically long due to the fact that the old connection is torn down only after the new connection has been established.

Since the roadside APs could evidently improve the wireless file delivery, especially when the mobility of vehicle and pedestrian can be predicted on familiar routes [25], my work attempts to provide a scheme of energy management on the availability and performance of digital devices under a realistic environment that the mobile devices could access the Internet source from cellular networks and Wi-Fi/White-Fi hotspots.

2.3.2 Mobility Management Protocols

Mobility management has become an important topic of research and various new protocols and standards have been designed to maximize the efficiency of mobility management, which mobile users can move around without noticing any change for the link service [90]. MIPv6 is designed to support the node mobility of IPv6 protocol, which MNs can move from one network to another and still maintain existing connections without changing their Home Address. There are three important entities, which are MN (Mobile Node), HA (Home Agent) and CN (Correspondent Node). Each MN is identified by two IP addresses: HoA (home address) and CoA (Care-of Address). The HoA is a permanent IP address obtained from its HA, which identifies the mobile node regardless of its location. MN is always reachable with its HoA; the CoA is changed at each new point of attachment and provides information about the MN's current situation. Whenever MNs move from the home link to a foreign link, they must be able to quickly detect its movement and acquire a new CoA. This CoA changes when MN moves to another foreign network. After confirming the uniqueness of new CoA at foreign network, the MNs sends send BU (Binding Updates) message to register this CoA with their HA and to notify CNs as needed. Once the BU has been acknowledged by CN, data packets can be transmitted through a path between the MN and CN directly [91] [92].

In [91], a practical testbed is established to examine the performance of MN vertical handoff in terms of the FTP handoff delay, which includes the different wireless radio access technologies such as WLAN, GPRS and TD-SCDMA. However, there are some drawbacks for the MIPv6 protocol. One is that periodic exchange of location update messages over the air interface consume available bandwidth; the other is that MNs' involvement in mobility management process consumes battery and incurs processing overheads [90].

PMIPv6 is a protocol that uses the same concepts as used in MIPv6, but modified to operate in the network part only instead of involving the MNs. PMIPv6 extends

MIPv6 signaling and reuses many concepts of MIPv6 such as HA. PMIPv6 benefits are evaluated regarding the increase of available bandwidth and TCP/UDP throughput over HSDPA and WLAN [90]. In [93], an implementation analysis of PMIPv6 is evaluated for the impact on the performance of PMIPv6 and reducing handover delays. In [94], the handover performance of PMIPv6 is compared with various representative existing host-based IP mobility management protocols, which handover delay of PMIPv6 is much lower than those of MIPv6 and HMIPv6. A handover delay analysis for mobile terminals with multiple interfaces based on PMIPv6 in [95] recommends to avoid excessive handover delays resulted from timer and configuration mechanisms in the IPv6 specification. With some performance optimizations, the handover latency of PMIPv6 could be around 100 ms, which is three times faster than MIPv6.

PMIPv6 extends MIPv6 signaling and reuses many concepts of MIPv6 such as the Home Agent (HA) functionality. The new principal functional entities introduced by PMIPv6 are LMA (Local Mobility Agent) and MAG (Mobile Access Gateway). LMA is the HA for the MN in a Proxy Mobile IPv6 domain.

- MAG is a function that typically runs on an access router. MAG is the entity that performs the mobility management on behalf of the MN. The first responsibility of the MAG is to detect the MN's movements on the access link and initiate the mobility-related signaling with the MN's LMA on behalf of the MN [96]. Furthermore, it establishes a tunnel with the LMA, and emulates the MN's home network on the access network for each MN.

- LMA is similar to the HA in MIPv6, but it has additionally required functional capabilities for supporting PMIPv6 [94]. It is responsible for maintaining the MN's reachability state while the MN moves around within a PMIPv6 domain. Every LMA must maintain a binding cache entry for each currently registered mobile node [96].

Once a MN enters a PMIPv6 domain and attaches to an access link, the protocol ensures that the MN is able to obtain its home address on any access network as long as it roams in the domain. According to the Per-MN-Prefix model, a unique home network prefix is assigned by the serving network to a MN for MN's exclusive use, and no other node shares an address from that prefix. As the prefix conceptually follows the MN wherever it goes within the PMIPv6 domain, given that the uniqueness of the global address in PMIPv6 is guaranteed by the per-MN prefix model, there is actually no need to perform the address configuration at the MN on the home address. If the MN can be prevented from going into an unnecessary address configuration, handover latency can be greatly reduced as compared to MIPv6. In addition, since MAG takes part in mobility-related signaling on behalf of the MNs, PMIPv6 can effectively reduce the overall handover delay by reducing the Proxy Binding Update round-trip time delay [95, 97].

A handover coordinator has been incorporated to enhance the PMIPv6 handover performance regarding handover delay and packet loss, where WiMAX and WiFi are interworking in a partially overlapping area [98]. An enhancement of PMIPv6 with simultaneous bindings for IP mobility is designed in [99], which can reduce the handover latency by proactively executing the configuration of the MN interface for the link to the target access router while the MN is still connected to the serving access router. In [97], a handover coordinator entity is utilized to operate in the overlapping region of interworking heterogeneous wireless networks (WLAN and WiMax networks) within a PMIPv6 domain. This HC ensures that packets are delivered to the MN as realtime as possible, thereby enhancing the handover performance by further reducing the handover delay and packet loss without incurring more signaling overhead in the air interface.

2.3.3 Cell Residence Time

The mobility model is of primary importance in wireless transmission analysis, which Cell Residence Time is defined as the time interval that mobile devices can access the service from Micro-cells. A systematic tracking for the random movement of mobile devices within a cellular environment is formulated mathematically in [100]. Based on this formulation, behavior of different mobility-related traffic parameters is characterized to indicate that the cell residence time both new and handover calls are well approximated by generalized gamma distribution. Moreover, depending on the street structure, the effect of changes in direction and speed of mobile users can be interpreted as contributing to an increase or decrease in the cell residence time. The effect of CRT distributions are quantified in [101] on cellular network performance, which the distribution of cell residence time depends on many factors such as cell shape, cell radius, user mobility and the paths a user follows. Based on the different CRT distribution such as Exponential, Erlang, Gamma, Uniform, Weibull and Deterministic, simulation results show that the call blocking probability is insensitive to the cell residence time distribution. Realistic mobility model is constructed in [102] to analyze different performance of call blocking probabilities based on exponential distribution and truncated Gaussian distribution for cell residence time with the same mean and standard deviation in the discrete event simulation. The cell residence time follows exponential distribution and Gamma distribution to explore HDT (handoff dwell time) effect in [103], while the cell residence times are independent identically distributed with nonlattice distribution in [104]. In [105], the probability distribution of CRT in the different quality zones is investigated assuming that the quality zones have specific shapes (i.e., circular-shaped cells). Experimental results indicate that the quality zone residence time tends to have a negative exponential distribution, according to specific distributions of speed and direction of movement of mobile users. In [106], based on a quite general mobility model, the new and the handover cell residence times can be approximated by a generalised gamma distribution. In [107], different distributions, such as negative-exponential, constant, Pareto, log-normal, gamma, hyper-Erlang, hyper-exponential and Coxian distribution, are utilized to model cell residence time.

2.4 Summary

Chapter 2 introduces the background and framework about Delay Tolerant Networks, Cognitive Networks, Spectrum Access, TV White Spectrum and Applications in vehicular networks. When a vehicle moves along a road segment, the factors that influence the wireless message transmission could be the channel conditions, distance to the serving BS, the speed of the vehicle, and the population of PU. The proliferation of delay tolerant applications (social networking updates, emails, updates over the air to mention just a few) can drive efficient utilization of available spectrum by SU under a CR enabled cellular network. As it is possible to reduce the energy cost by delaying the wireless transmission until high-rate Wi-Fi/White-Fi hotspots become available, significant research efforts have been placed on energy efficient data transmission that takes into account the delay-sensitivity of mobile applications with delay deadlines. Especially the problem of energy efficient transmission under delay constraints over wireless networks has been studied in section 2.1.

In the addition to the above, as the vacant TVWS spectrum are permitted to be used in several countries, the new available spectrum will unfold new possibilities in data transmission with strong potential to further decrease energy consumption. According to the historical information about availability of primary channels, PU traffic detection and load balancing are two important issues for channel selection algorithms for SUs. Hence, queuing theory and its applications have been described in subsection 2.2.1 to analyze the opportunities for wireless transmission. Then, there has been many research focusing on spectrum decision strategies for SU transmission, which can be divided into two classes in subsection 2.2.2, namely sensing-based scheme and probability-based scheme.

Finally, since the ubiquitous deployment of wireless Internet connectivity has brought growing pressure on cellular data networks, many research have suggested that the architecture of next generation wireless networks would constist of Macro-cells (BS) and Micro-cells (hotspots) to offload wireless traffic. Since micro-cells could be installed in addition to the macro BSs, an elementary strategy that can increase the capacity of a cellular network and reduce the usage of cellular data networks is to utilize the road-side high-speed hotspots, which is the so-called heterogeneous network. To propose a scheme in which mobile terminals can intelligently switch among cellular, Wi-Fi and White-Fi interfaces, more work on the characteristics of heterogeneous network has been described in section 2.3. Following the overview of simultaneous use of multiple interfaces on mobile devices, various studies on mobility management protocols have been discussed in subsection 2.3.2. These protocols are designed to support the mobility of mobile devices, where handover delay and packet loss are two key indicators to evaluate their performance when they are inplemented in the mobile embeded systems. Moreover, in order to appropriately characterize the service time within a coverage of Micro-cells, an accurate mobility model of mobile user's residence time in the Macro-cells is required. Cell residence time are assumed to follow a particular distribution in previous research described in subsection 2.3.3, which are related to mobility attributes, such as user velocity, user direction and route selection of mobile users.

Chapter 3

Multiple Interfaces for Wireless Transmission

In Chapter 3, the framework of cellular networks with several high-speed wireless hotspots is proposed and several relevant standardization activities are explained. As a Secondary User, the mobile users should effectively avoid the Primary connections and lower itself energy cost for wireless transmission. Since it becomes feasible to estimate the PU connections by contacting a trusted database containing the information of PU traffic, a scheme is proposed that explicitly utilize the distribution of SU traffic loads to provide load-balancing. By modeling the problem under investigation as an M/M/K/L queuing system introduced in Appendix A, the performance of the SUs competition is analyzed under various traffic blocking thresholds and queuing delays. To this end, the optimal number of SU traffic capability can be estimated to be served simultaneously in the queuing systems as well as optimizing the frequency channel utilization. A wide set of numerical investigations reveal how data transmission delays and the use of available white spaces can reduce the energy consumption without affecting PU traffic.

3.1 Wireless Accesses

In real terms, handover among Wi-Fi networks (White-Fi channels) and cellular networks incurs the interface switch energy cost and connection establishment energy cost.

Wireless channel conditions are also highly correlated to the geographical location of the nodes (i.e., path loss), therefore accurate wireless networks condition estimation is highly critical in cognitive networks for the efficient usage of the available spectrum at a given location and session duration. In our model, we calculate the theoretical gain of energy savings by a judicious use of multiple wireless interfaces.

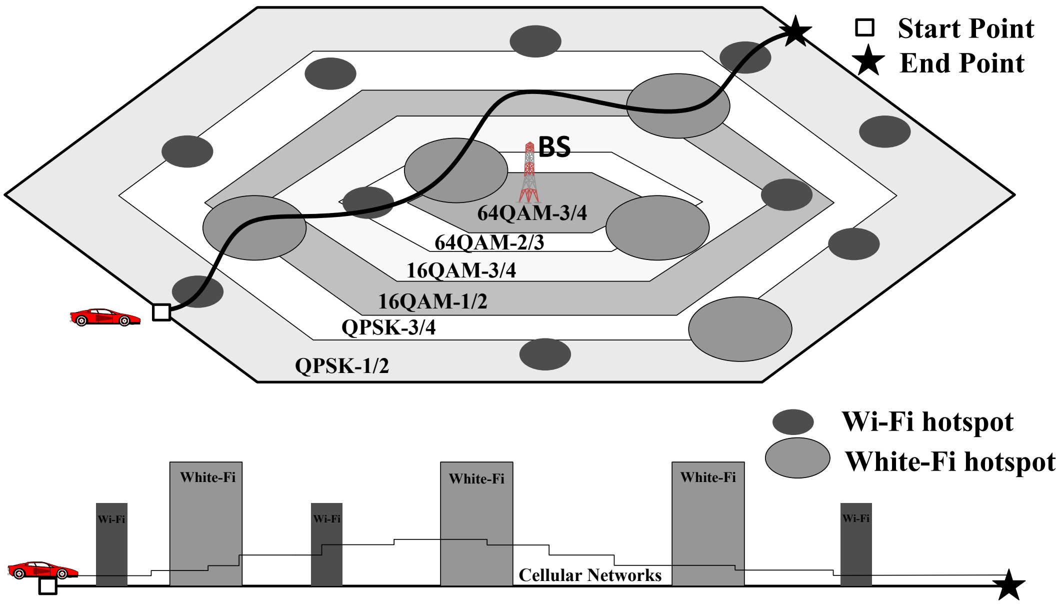


FIGURE 3.1: The bit rate of mobile users within the coverage of the Base Station (BS) ring range

3.1.1 Deployment of Cellular Networks

The ability to perform Adaptive Modulation and Coding (AMC) based on received signal quality indicators allows the cellular networks to dynamically use higher-order modulation, up to $64QAM$, within locations with increased signal to noise ratio and lower-order modulation such as $QPSK$ [108] [109]. To this end, the cell can be separated into n concentric rings of radii $R_i, i = 1, \cdots, N$ according to the distance between the wireless device and its serving BS as shown in Figure 3.1. The set of available modulation and coding schemes are denoted by $\{M_{R_1}, M_{R_2}, \cdots, M_{R_N}\}$. Consequently, each circular region with distance R_i to the BS corresponds to a different constellation size and coding scheme. For example, 64QAM-3/4 means the modulation and coding scheme is 64 with 3/4 coding rate, while QPSK-1/2 denotes the modulation and coding scheme is 4 with 1/2 coding rate.

Let r denote the coding rate, and d_{BS} be the (Euclidean) distance between the mobile node and the serving BS. Then, since the spectral efficiency is dominated by the distance between mobile user and the BS, the spectral efficiency (bits/s/Hz) $IEC(d_{BS})$ is given by [108]:

$$IEC(d_{BS}) = r \cdot \log_2(M_{d_{BS}}) \;\; bits/s/Hz \tag{3.1}$$

Concerning the energy consumption in data transmission, it can be identified by the following main sources: (1) the electronic circuit at the nodes of SU; (2) the energy dissipated to transmit the data; and (3) the energy consumed to receive corresponding data. The energy consumption regarding SU wireless data transmission can be formulated as in [110], shown in the equation below where e_d represents the energy consumption of the Operational Amplifier, B denotes the data transmission rate, and τ is the transmission time.

$$f(d_{BS}) = (e_{rx} + e_{tx} + e_d \cdot d_{BS}^{\eta})B \cdot \tau \tag{3.2}$$

$$= \begin{cases} (e_{rx} + e_{tx} + e_{los} \cdot r_{BS}^2) \cdot B \cdot \tau & if \; d_{BS} \leq d_{ths} \\ (e_{rx} + e_{tx} + e_{mp} \cdot r_{BS}^4) \cdot B \cdot \tau & if \; d_{BS} > d_{ths} \end{cases} \tag{3.3}$$

where e_{rx} is the Receiver electronics consumption and e_{tx} is the Transmitter electronics consumption, and e_{los} denotes the one bit energy consumption with the free space (line-of-sight) propagation losses (the distance to the BS $d_{BS} \leq d_{ths}$) and e_{mp} is the energy consumed per bit in transmission in the longer-distance case that the signal is attenuated by multi-path fading (Two-ray Reflection). Let P_{los} and P_{mp} denote the transmitted signal powers in these two cases respectively [111], we obtain

$$P_{los} = \frac{P_{rx}(4\pi)^2}{\lambda_{wl}^2 G_{tx} G_{rx}} \cdot r_{BS}^2 \tag{3.4}$$

$$P_{mp} = \frac{P_{rx}}{(h_{tx}h_{rx})^2 G_{tx} G_{rx}} \cdot r_{BS}^4 \tag{3.5}$$

where λ_{wl} is the wavelength, h_{tx}/h_{rx} is the transmit/receive antenna height, and G_{tx}/G_{rx}

is the transmitter/receiver antenna gain. Therefore, the energy consumption for transmission is given by:

$$f(d_{BS}, t) = (e_{rx} + e_{tx})F + \sum_{t=\tau_k}^{\tau_{k+m}} P_{los}(or\ P_{mp}) \cdot t \tag{3.6}$$

where τ_{k+m} is the time slot that finish the data transmission.

We adopt an LTE system with a 20 MHz bandwidth throughout the investigations hereafter. In LTE channels, the average sector throughput could be utilized to quantify the total capacity throughout the sector or site coverage. With the aggregate of the data rates and the number of concurrency users within the sector, we can determine the network throughput for individual mobile user. Let $\Gamma \in \{\tau_0, \tau_1, \tau_2, \cdots, \tau_T\}$ denote the time slots along the whole route of mobile users. During the time duration $\{\tau_k, \tau_{k+1}, \cdots, \tau_{k+m}\}$ from all the time slots above, the mobile user could transmit a portion of the file to the BS. Hence the energy cost of the whole file transfer via the cellular interface can be calculated as follows,

$$E_{cell} = \Sigma_{t=\tau_0}^{\tau_T} P_{elec}(t) \cdot t + \Sigma_{t=\tau_k}^{\tau_{k+m}} e_{tx}(IEC, d_{BS}) \tag{3.7}$$

Where P_{elec} in the above expression denotes the power consumption of the cellular interface electronics and $e_{tx}(IEC, d_{BS})$ represents the energy cost for the op-amp, which is related to the distance between the mobile user and the BS and the spectral efficiency.

Within cognitive radio (CR) networks, the SU transmission is trying to explore white space of wireless spectrum while at the same time avoiding to interfere with the PU traffic. For the purpose of avoiding PU transmission, as a SU, the fundamental assumption is that wireless nodes in vehicles use spectrum sensing or query the database which maintains information about the available channels for the details of the local radio environment.

3.1.2 Modelling Wi-Fi Operation

In general, the Wi-Fi transceiver has four basic modes of operation: Sleep, Idle, Transmit (TX), and Receive (RX). From the report in [112], the power consumption of wireless

NICs is not decreasing with the new generation products, resulting from the fact that more power consuming subcomponents were used in Wi-Fi systems. The report likewise brings out the fact that power consumption is in proportion to the applied voltage of the Wi-Fi module. Moreover, we can see that the NIC power consumption in Idle state is comparable to their TX&RX power. This is because in Idle mode, the Wi-Fi module has to listen frequency channels and wait for incoming packets, leading to the fact that both the analog circuits and Analog-to-Digital Converter (ADC) operate at full workload as in the RX mode. On the contrary, it is the sleep mode actually that can introduce substantial power savings [113].

In current Wi-Fi chips that target smartphones, mobile gaming and portable consumer electronic devices utilize the SDIO (Secure Digital Input Output) and GSPI (Generic Serial Peripheral Interface) interface for the design of low power embedded systems. Although the above studies in [112] [113] just present the figures of WLAN products including the PCI, mini-PCI, CardBus, PCMCIA interface and so on, and the power consumption of Wi-Fi chips in smartphones and tablet PCs should differ with the above studies shown, they should dissipate consistent relative power.

We assume that there are several Wi-Fi hotspots within the coverage of a macro BS. Within the domain of Wi-Fi hotspots, mobile users are always associated with a serving AP, and we further assume that APs are randomly located within the cell coverage area [114]. The set of AP locations could be denoted by $\Phi_\alpha = \{x_1, x_2, \cdots, x_k\}$. The research in [115] depicts three distinct phases for TCP connectivity, where the entry phase and exit phase provide very weak connectivity with higher loss rates and delay. In contrast during the production phase, a connectivity window provides a significant higher throughput. Depending on the signal strength received from an AP by a mobile user and the user's speed, it will maintain a constant bit rate in the production phase of the connection between the user and an AP. Hence if the user's speed does not change too much, the effective bit rate can basically keep fixed for the specified user [116]. Therefore, without loss of generality, we assume that the associated APs provide a constant data rate to the mobile user within their coverage.

3.1.3 Spectrum Management Models

As mentioned above, wireless channel conditions are highly correlated to the actual geographical location of the mobile users since the path loss is a function of the distance to the AP. In this case, accurate wireless networks condition estimation is critical. Meanwhile, for the purpose of avoiding PU transmission, the SU nodes could utilize spectrum sensing or query the database which maintains information about the available channels for the details of the local radio environment.

One approach to utilize the white space spectrum at a given time and location is to verify with a database available channels. From historical statistics, SU could verify an average Wi-Fi (White-Fi) availability within a 24 hours period, thereby quantifying the average coverage of White-Fi in this domain at a given time. On the other hand, if such a trusted database is not available in this area, the mobile devices have to sense the spectrum.

In the sensing case, when we consider the optimized overall energy saving schemes, the sensing energy cost should be factored as essential part. Energy consumption per sensing cycle E_{scl} is given by the sum of the total energy to scan all the assigned channel E_{scan}, and the total energy to switch between these channels E_{sw}. We assume the number of channels for White-Fi is C, hence the mobile device has to switch $(C-1)$ times when hopping from one channel to another [117].

$$E_{scl} = C \cdot E_{scan} + (C-1) \cdot E_{sw} \tag{3.8}$$

Apart from that, the locations of Wi-Fi AP based White-Fi is another crucial factor in the wireless transmission efficiency. If the information of AP locations is not available for mobile devices, we have to predict the PUs actions, the probabilities of establish a connection, and service duration according to the density of White-Fi in this domain.

3.1.3.1 Optimal Transmission Time of SUs

In the proposed model as detailed hereafter, we assume that wireless nodes gather PU traffic information from a historical database in order to predict the traffic pattern of

PUs over a short-term period. In the database server, there are two types of information about primary channels. One is the 24-hours traffic characteristics of different channels in several months [32], the other is the noise power level in different channels that updated by SU devices constantly. Based on the above framework, a SU has to send a query to the database server for an available channel to transmit. According to the available channel information at the same time slots in previous days and long-term statistics regarding channel availabilities, the database server will certificate the noise power levels by comparing to a threshold. Then the best candidate channels for the inquired SU will be determined.

As there is significant previous research regarding the prediction of PU connections arrival rate and holding times, we aim to utilize such results and design an algorithm for SU connections according to the periodicity of PU traffic pattern. In this scenario, we consider the time horizon for the estimation to be relatively long term (such as for example in terms of hours) instead of short term (fraction of a second), that is, the prediction of PU traffic is stable and unchangeable in a relative long-term time. Therefore, under the assumption of the probability-based channel selection scheme [1], the SU will select its operating channel from all of the M candidate channels to achieve some form of load-balancing based on the stochastic traffic prediction, such as the arrival rates and service times. The packet arrival processes of the PU connections are modelled as Poisson. Applying the Probability-based Scheme theory [1], the average system time S is given by,

$$E[S] = E[W] + E[T] \tag{3.9}$$

where W denotes the waiting (queuing) time and T presents the extended data transmission time. Then, let $\tilde{X}_p^{(k)}$ and $\tilde{X}_s^{(k)}$ be the extended service time of PU and SU respectively due to the missed detection and false alarm probabilities. $\lambda_p^{(k)}$ and λ_s are the average arrival rates of the PU connections at channel k and the SU connections. $\rho_p^{(k)}$ and $\rho_s^{(k)}$ are the busy probabilities resulting from the PU and SU connections at channel k.

$$E[T] = \sum_{k=1}^{K} p_{pb}^{(k)} \frac{E[\tilde{X}_s^{(k)}]}{1-\rho_p^{(k)}} \tag{3.10}$$

$$E[W] = \sum_{k=1}^{K} p_{pb}^{(k)} \frac{\frac{1}{2}\lambda_p^{(k)} E[(\tilde{X}_p^{(k)})^2] + \frac{1}{2} p_{pb}^{(k)} \lambda_s E[(\tilde{X}_s^{(k)})^2]}{(1-\rho_p^{(k)})(1-\rho_p^{(k)}-\rho_s^{(k)})} \tag{3.11}$$

$$\rho_p^{(k)} = \lambda_p^{(k)} E[\tilde{X}_p^{(k)}] \tag{3.12}$$

$$\rho_s^{(k)} = \lambda_s^{(k)} E[\tilde{X}_s^{(k)}] \tag{3.13}$$

where the vector $p_{pb} = (p^{(1)}, p^{(2)}, \cdots, p^{(k)}, \cdots, p^{(K)})$ presents the set of distribution probabilities for selecting each candidate channel $k \in (1, K)$. We utilize probability-based scheme to find the optimal probability vector p_{pb} in order to balance the PU and SU traffic load among multiple channels with the constraint that $p^{(1)} + p^{(2)} + \cdots + p^{(K)} = 1$. Moreover, this optimal p_{pb} vector can maximize the transmission time in each time slot in order to accommodate more SU traffic. On the basis of this optimal vector, the SUs could directly access to the preselected channel by estimating the PU traffic characteristics from historical stored data.

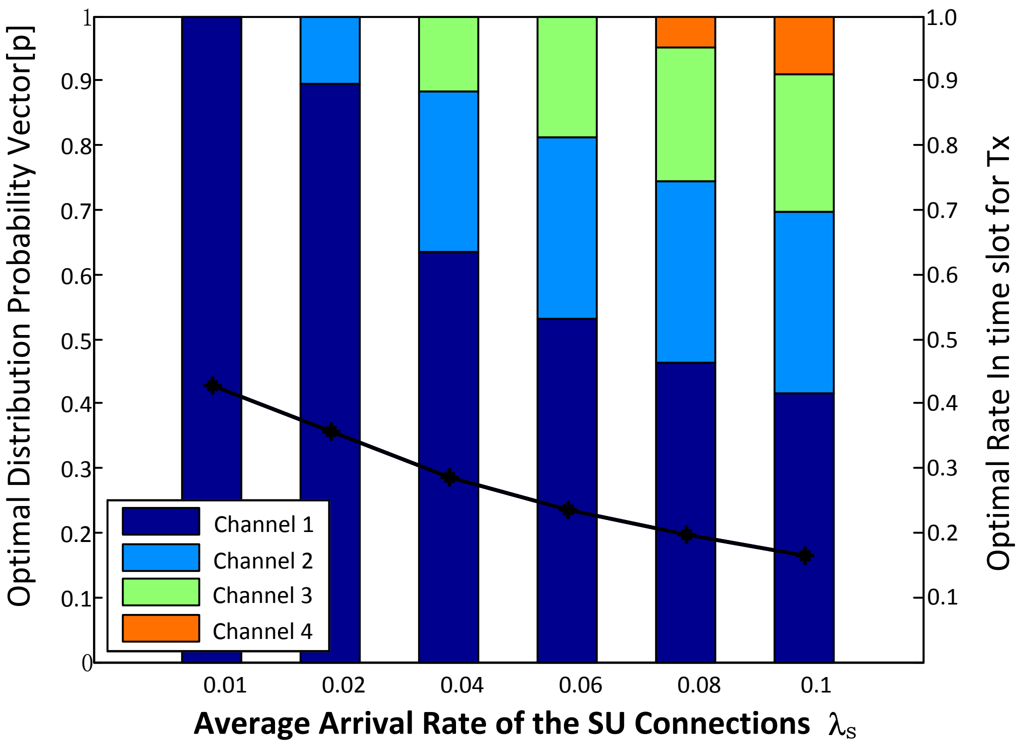


FIGURE 3.2: SU traffic pattern with different average arrival rates

According to a four-channel system with the traffic parameters in Table 3.1, which assumes that the data rate of the primary and the secondary connections (bits/slot) are equal to 1, the results in [1] have been reproduced as shown in Figure 3.2. It displays that when the SU traffic is relatively light, i.e., the arrival rate being 0.01, the SU traffic prefers to select channel 1 as their transmission frequency channel which has the lightest PU traffic. However, when the SU arrival rate becomes higher, some secondary connections tend to choose other candidate channels for data transmission. Based on the optimal distribution probability vector, the same service time for secondary connections and minimal overall system time, we can have the optimal time duration rate as presented on the right y-axis of Figure 3.2 for SU traffic in one time slot. In our model, it is assumed that the average arrival rates of the SU connections are under a threshold which restricts the SU traffic for better QoS service.

TABLE 3.1: Parameters of PU traffic, where $E[X_s] = 10$, $\lambda_s = 0.06$ [1].

Num of channel	arrival rate (λ_p)	service time (bits/arrival) ($E[X_p]$)	probability of SU traffic (p_{opt})
1	0.01	20	0.5308
2	0.01	30	0.2828
3	0.02	20	0.1864
4	0.02	35	0.0000

Recall that the cell has been separated into N concentric rings as shown graphically in Figure 3.1. When a SU message needs to be transmitted, firstly, it will be placed into a specialized queue for corresponding ring. Secondly, the system of queue is sensing for the available vacant frequency channels that the PU connections do not occupy and tries to distribute the message from SU queue into the candidate channels. We assume without loss of generality that all SU messages have the same data length. As a result, the SU connections in all of the frequency channels are assumed to have the same service time. From the above discussion, it is possible to estimate the optimal time duration for SU traffic in a specific time slot. Then the problem can be simplified into a situation that only SU messages compete for all of the frequency channels, i.e., without competing with the PU connections. Under the above defined framework, the central focus in the sequel is to calculate the number of SU that can be accepted in the different co-eccentric rings in the network under the constraints of service delay elasticity and the blocking probability.

3.1.3.2 Optimal Number of SUs

In this section, we turn our focus on the SU message competition without competing PU connections. Figure 3.3 gives an example of the physical queues for the case of K frequency channels and N concentric rings with different modulation and coding schemes. When the traffic of the SUs needs to be transmitted in the system, it can be inputted to the queue for the SU connections. This proposed channel selection model could approximate the virtual SU message queue using an M/M/K/L queuing system, as described in Appendix A. If the number of SUs is large, the input traffic of the virtual queue can be modeled as a Poisson process, where K is the number of servers and L is the finite number of waiting positions for each queue.

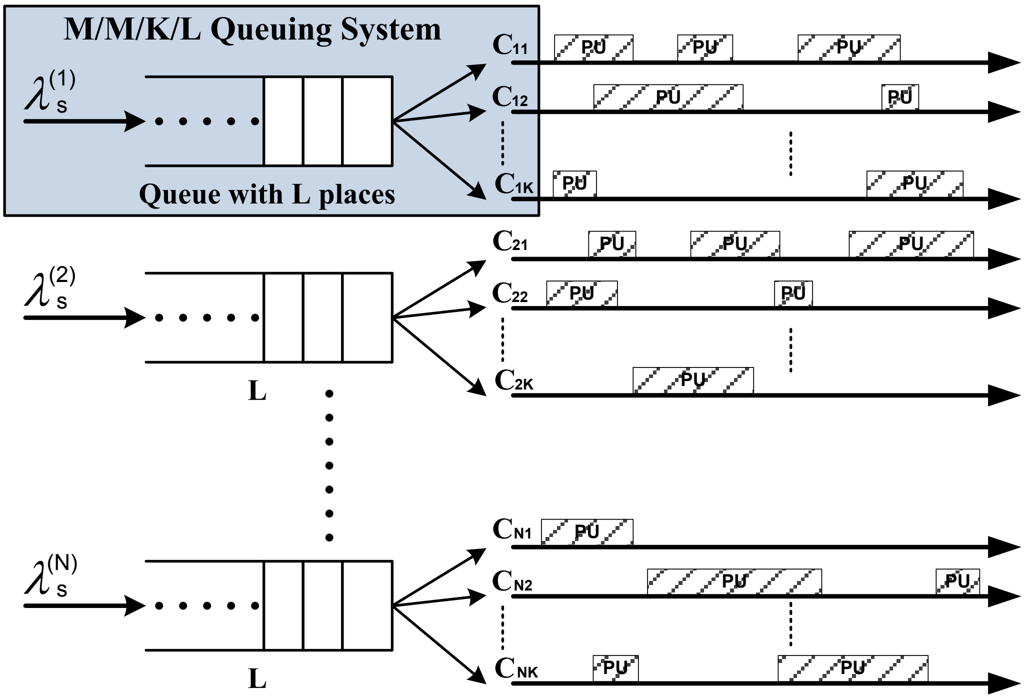


FIGURE 3.3: Access for SUs modeled as an M/M/K/L queuing system

Let $\lambda_s^{(i)}$ denote the average number of the SUs per unit time in ith concentric ring of radii R_i and L denote the number of unit time as a kind of queue length. Therefore, each queue of this system, namely each ring, can accommodate $\lambda_s^{(i)} L$ number of the SUs. Given the set of candidate channels $\Omega = \{1, 2, \cdots, K\}$ and the set of concentric rings $\Re = \{1, 2, \cdots, N\}$, we denote C_{ij} to be the capability of the SUs in ring $i \in \Re$ within the channel $j \in \Omega$ and have

$$C_{ij} = IEC(d_{BS}) \cdot \frac{B}{K \cdot N} \quad (bit/s) \tag{3.14}$$

where F is the size of SU message and B represents the bandwidth available at the BS. Let μ_{ij} represent the service rate of SU connections using the frequency channel j in ith ring, we have,

$$\mu_{ij} = \frac{C_{ij}}{F} \tag{3.15}$$

Let ρ_i denote the occupation rate (offered traffic load), we have

$$\rho_i = \frac{\lambda_s^{(i)}}{K \cdot \mu_{ij}} = \frac{F \cdot \lambda_s^{(i)}}{K \cdot C_{ij}} \tag{3.16}$$

Assume that p_m is the probability that there are m SU message in the system and p_{thres} is the blocked traffic rate threshold, therefore we have

$$p_m = \begin{cases} \frac{\rho_i^m}{m!} \cdot p_0 & m \leq K \\ \\ \frac{\rho_i^K}{K!} (\frac{\rho_i}{K})^{m-K} \cdot p_0 & K < m \leq \lambda_s^{(i)} L + K \end{cases} \tag{3.17}$$

subject to:

$$\sum_{m=0}^{\lambda_s^{(i)} L+K} p_m = 1, \tag{3.18}$$

$$p_{(\lambda_s^{(i)} L+K)} \leq p_{thres} \tag{3.19}$$

Note that the constraint of (Equation 3.19) depicts that the blocking rate of SU connections in the virtual queue should be lower than the predetermined threshold p_{thres}.

3.2 Wireless Interface Switch

Conventional Wi-Fi is deployed in the unlicensed 2.4GHz spectrum (IEEE802.11 b/g/n), while several wireless card vendors consider to trigger a new amendment in order to add Wi-Fi-like TVWS technology into the already existing IEEE 802.11 standards body. Moreover, the lower transmission frequency (in the range 400 800 MHz) leads to the increased coverage of Wi-Fi-like AP in the TV band. In order to identify the effect of occupying frequency channels in the interference under various use cases, cognitive devices have to seek proper approaches to provide reliable communications without affecting PUs.

In the simulations hereafter, we assume that APs (Wi-Fi, White-Fi) are uniformly distributed over a typical urban environment. For instance, we consider 20%, 50%, and 80% of coverage from such small cells in order to evaluate spectral availability under scenarios where Wi-Fi/White-Fi coverage is low, medium and high. Furthermore, the alternative wireless networks (Wi-Fi, White-Fi) may be available at limited locations or time durations. Therefore, the challenge is whether a mobile device that will transfer N MB of data should search for alternative networks to transmit the data and possibly delay the transmission in order to achieve minimal energy cost.

Let E_{TVav} be the energy cost for data transmission as the White-Fi radio access is available for mobile users, which is including the transmission cost via White-Fi radio E_{txTV} and sensing cycle cost E_{scl}. When the White-Fi access is not available, it is assumed that there is a energy cost E_{TVun} when the radio circuit has been active.

$$E_{TVav} = E_{txTV} + E_{scl} \tag{3.20}$$

$$E_{TVun} = E_{scl} \tag{3.21}$$

Meanwhile, based on historical statistics from roadside APs in urban environment, it is assumed that mobile users can directly associate the roadside APs on a specific spot which could evidently reduce the sensing cost on the Wi-Fi radio circuits. Since the location awareness of the different hotspots could be not accurately enough it is expected that when the mobile user is not within the coverage of Wi-Fi hotspots, there is still some energy consumed by the circuit of Wi-Fi interface in the mobile device. Therefore, the Wi-Fi access cost E_{Wav} and the cost that Wi-Fi is not available can be denoted by,

$$E_{WFav} = E_{txWF} + E_{assoWF} \tag{3.22}$$

$$E_{WFun} = E_{cirWF} \tag{3.23}$$

In the above expression E_{txWF} expresses the energy cost through the entire wireless transmission, E_{assoWF} is the energy consumption for Wi-Fi AP association which remains a small proportion from overall Wi-Fi circuit cost, and E_{cirWF} is the only Wi-Fi circuit cost. Since the interface for the LTE (primary network) is always on, there is no energy cost for connection establishment. Therefore, the energy cost to transfer the data via the LTE interface is E_{LTE}.

Let P_{TV} denote the probability that a White-Fi network is available. The expected energy cost of attempting to use the White-Fi network is as follows,

$$E_{TV} = P_{TV} \cdot E_{TVav}(L_{TV}) + (1 - P_{TV}) \cdot E_{TVun} \tag{3.24}$$

Similarly, the energy cost of attempting to use Wi-Fi access is,

$$E_{WF} = P_{WF} \cdot E_{WFav}(L_{WF}) + (1 - P_{WF}) \cdot E_{WFun} \tag{3.25}$$

where L_{TV} and L_{WF} represent the file size transmitted within one White-Fi and Wi-Fi hotspot respectively, while L_{total} is the entire file size of data and P_{WF} is the probability that Wi-Fi access is available.

In the case of multiple alternative interfaces, the system should choose the network with the most expected energy saving. In this work, the primary network is cellular networks and the alternative networks are White-Fi and Wi-Fi. If the data size transmitted by White-Fi L_{TV} or by Wi-Fi L_{WF} is smaller than the entire data size L_{total}, which means the entire data cannot be transmitted within the current cell, the system has to determine whether to delay the transmission to the next wireless hotspot cell. Otherwise, it has to utilize LTE access to transmit the remaining data portion. Meanwhile, if all of the data could be transmitted within the high-speed access coverage like Wi-Fi and White-Fi, it will not consider the handover for the remaining data. The energy cost for multiple interfaces can be denoted by,

$$
\begin{aligned}
E_{multi3} &= P_{TV} \cdot E_{TVav}(L_{TV}) + (P_{WF} - P_{WT}) \cdot E_{WFav}(L_{WF}) \\
&+ (1 - P_{TV} - P_{WF} + P_{WT}) \cdot E_{LTE}(L_{LTE}) \\
&+ (1 - P_{TV}) \cdot E_{TVun} + (1 - P_{WF} + P_{WT}) \cdot E_{WFun} \\
L_{total} &= L_{TV} + L_{WF} + L_{LTE}
\end{aligned} \tag{3.26}
$$

where P_{WT} is the probability that both Wi-Fi and White-Fi access are available. And it is assumed when White-Fi and Wi-Fi access are both available, White-Fi access has a higher priority than Wi-Fi network. Furthermore, we categorize the wireless transmission strategies for mobile applications into several different categories as shown in Figure 3.4.

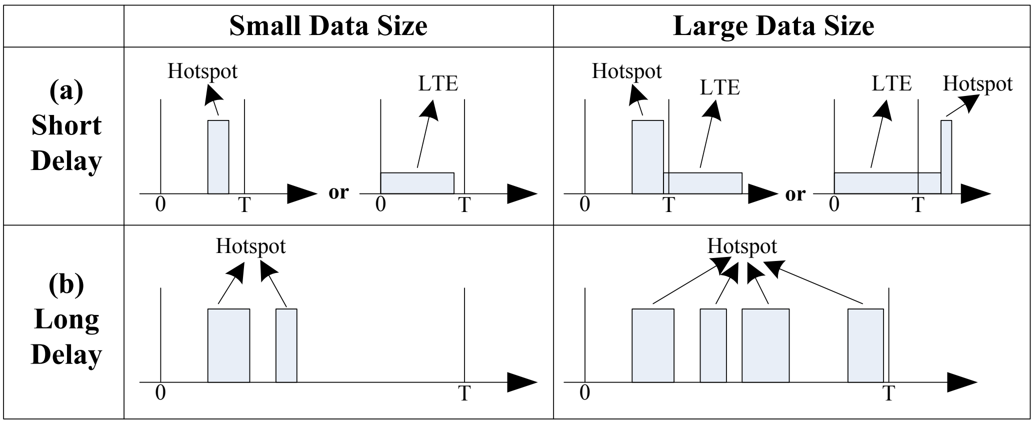


FIGURE 3.4: Transmission strategy for different data size under delay constraints

1) Large size files with short delay constraint (YouTube like applications, 10 seconds delay constraint): (i) In the only-LTE-coverage area, the system will predict whether the mobile user can move into the next hotspot under the delay constrains of the applications. If this is not possible, the user has to utilize LTE interface to transmit a proportion of data file in order to meet the applications' requirements before moving into next hotspot; (ii) Within the Wi-Fi and White-Fi coverage area, if all the data cannot be transmitted, the system has to determine whether delay the transmission to next hotspot or immediately transmit the remaining data via LTE. In order to provide better user experience, the system might have to start the wireless transmission immediately to

secure enough downloaded data in local storages for video playback rather than waiting until next high-rate hotspot.

2) Small size files with short delay constraint (podcast, audio files): the system will launch the file transfer immediately, no matter what kind of wireless radios it could access for transmission.

3) Delay-tolerant applications (email, social network and APPs updating, 1 min delay constraint): if the mobile applications are delay tolerant, the mobile users will have enough time to moves into the next wireless hotspot, such as White-Fi and Wi-Fi. Therefore, the wireless transmission could be always executed via the high throughput interfaces.

3.3 Bundle Protocol for DTNs

In 2003, Intel has implemented a prototype DTN architecture under the Linux operating system, in which DTN (Bundle) gateways are in charge of buffering messages in nonvolatile storage devices. It is well noted that the DTN design can be overlaid upon the TCP/IP based Internet easily [48]. From then on, it is assumed that all of the DTN system should be an entity with a bundle layer. The first use of Bundle Protocol for DTN is in 2008 [118]. In these tests, a bundle node transferred images from a Low Earth Satellite (LEO) belonging to the UK Disaster Monitoring Constellation (UK-DMC) to provide store-and-forward service in DTN networks via Bundle Protocol. The Bundle Protocol (BP) is defined in RFC 5050, while the DTN architecture introduces a bundle layer in RFC 4838. Bundles can be stored in an intermediate node for an excessive amount of time (minutes, hours, or even days) [119]. By using Bundle Protocol as a store-and-forward protocol and LTP (Licklider Transmission Protocol) as a convergence layer protocol, the DTN transport layer service has been evaluated in [120].

In a DTN, the bundle layer is placed below the application layer and hides the actual communication layers [49] [118] [119] as presented in Figure 3.5. The Bundle Protocol relies on the services of a convergence layer that is underneath the bundle layer. The Bundle Protocol allows application programs to communicate across the lower-lower protocols under conditions that involve long network delays or disruptions [121]. The

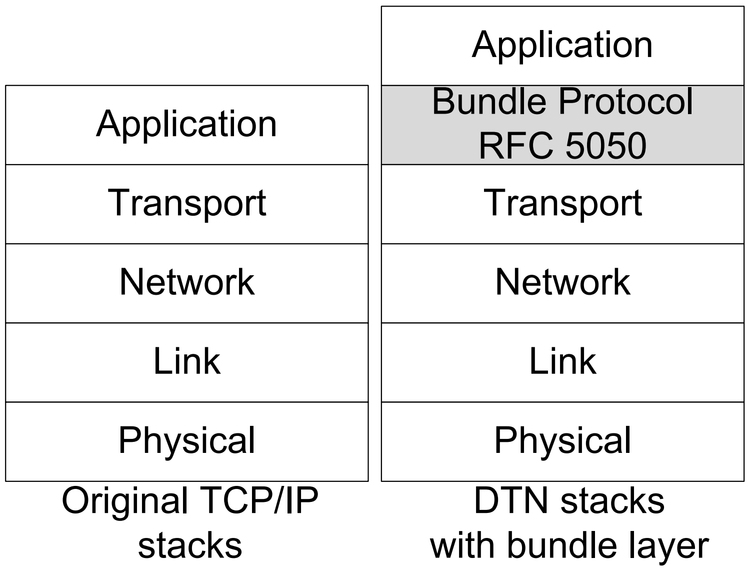


FIGURE 3.5: The architecture of DTN layer

convergence layer bridges Bundle messages and link layer in order to practically make the network communication. In order to deploy DTN in practice, the implementation in [49] utilized TCP/IP as link layer protocol based on Android mobile phones.

3.4 Performance Investigation

The related parameters used in general scenarios where a vehicle is moving towards the BS are summarized in Table 3.2 and we use the model in [110] to calculate the energy consumption of RF module and electronic circuits. The results of energy consumption were obtained via MATLAB based simulation by focusing on the uplink throughput to serving BS and it is assumed that the radius of the cellular BS to be 1000 meters.

TABLE 3.2: System Parameters

Symbol	Definition	Value
F	Message Size	50 Mbits
P_{rx}	Received power threshold	-52 dBm
v_a	Vehicle average speed	8 m/s
BW	Bandwidth	10 MHz
e_{tx}	Transmitter electronics consumption	50×10^{-9} J/bit
e_{rx}	Receiver electronics consumption	50×10^{-9} J/bit
V_{DD}	Power supply	1.8 V

We provide an example of optimal allocation of the PU connections and the blocking rate probabilities of SU connections in Table 3.1. For simplicity, we assume the channel number $K = 4$. Note that in Table 3.1, the arrival rate and service time of the PU connections in all of the K channels are not identical. In addition, as shown in the results of optimal distribution probability vector for SU transmission, we can see that increased probabilities of SU connections are allocated to the channels whose traffic load is expected to be less. As a result, the wireless node on the vehicle would always try to select the optimal channel for data transmission and avoid the interruption on the PU transmission. Based on the above discussion and the parameters from Table 3.1, the optimal distribution probability vector for transmission time of vehicle p_{opt} is $(0.5308, 0.2828, 0.1864, 0.0000)$.

Recall that the SUs must monitor all frequency channels to sense the arrival of PU connections. In this model, we assume that the PU traffic load is stable within a short time duration. Therefore, from now on, we use queuing system modeling to analyze the traffic of SU connections without consideration of PU connections. The SU virtual queue senses the frequency channel in an increasing order, from 1th to the Mth channel. When initiating a data transmission, the queues first check the availability of frequency channel 1, then $2, 3, 4, ..., M$. If all of the frequency channels are unavailable, the transmission is blocked. Here, for the SU virtual queue, the availability of one frequency channel means that the channel is not occupied by the PU connections.

TABLE 3.3: Simulation Results ($p_{thres} = 0.10$)

Num of Ring	Num of time unit	Max λ_s	Packet length F (Mbits)	Blocking probability
1	42	7	1	0.0490
2	23	11	1	0.0909
3	84	14	1	0.0477
4	77	22	1	0.0909
5	167	29	1	0.0805
6	188	33	1	0.0909
total	581	116	1	N/A
optimal	188	199	1	0.0955

Table 3.3 presents the simulation parameters and results with a blocking traffic threshold $p_{thres} = 0.10$ for the SU queue. It can be clearly seen from the results that when the mobile nodes are moving towards the service BS, the length of time units in each ring

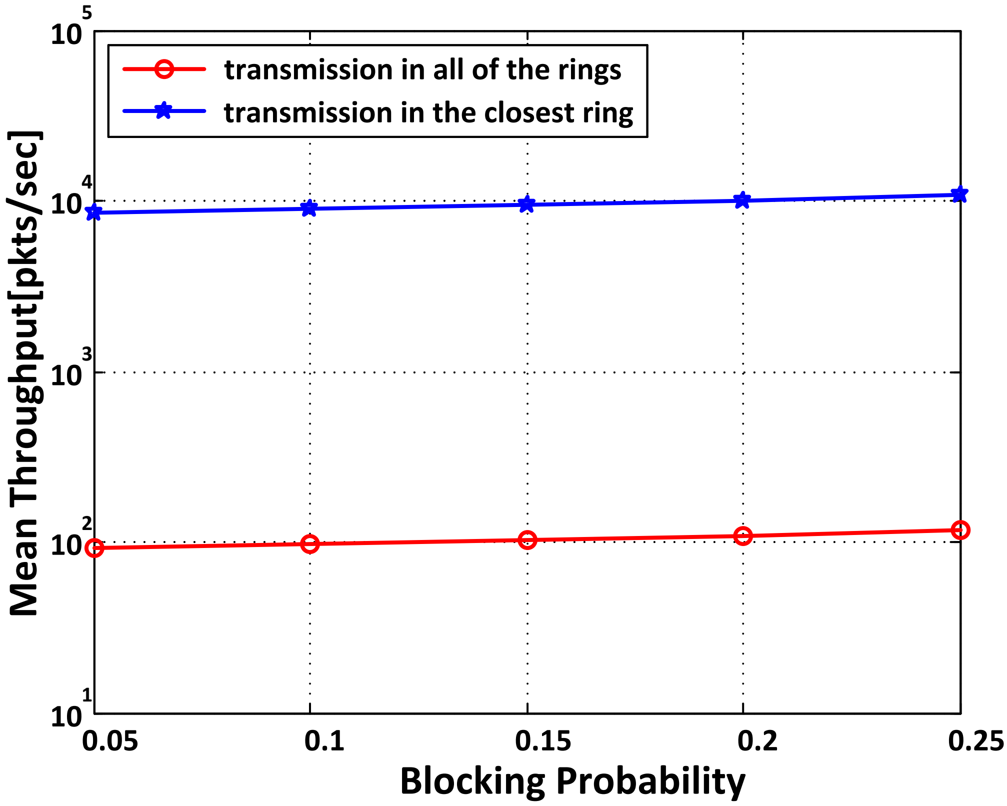


FIGURE 3.6: mean throughput of the SUs vs. different blocking probability threshold

becomes less owing to the increasing capability of rings in terms of transmission rate. Moreover, with the decreasing distance between wireless nodes and the BS, the ring could accommodate a larger number of SU messages, that is, the SU virtual queue could accommodate a larger number of the SUs simultaneously.

In this model, we assume that there are two situation for a comparison in order to show the benefit of energy consumption: 1) the bandwidth of the service BS is divided into six equal parts, and each ring could only use one part for transmission; 2) the bandwidth is occupied exclusively by the last ring which is closest to the service BS. For instance, in the 7th line of Table 3.3, the SUs buffer all the message and move into the area of the ring which is closest to the service BS. In this case, the SUs in the last ring occupy all of the bandwidth for message transmission, that the capacity and throughput will be the maximum possible since this ring can support the higher constellation. From Figure 3.6 we can observe that, if the message carried by the SUs is highly delay-tolerant, when the SUs transmit message in the closest ring of the BS, the throughput could be significantly improved.

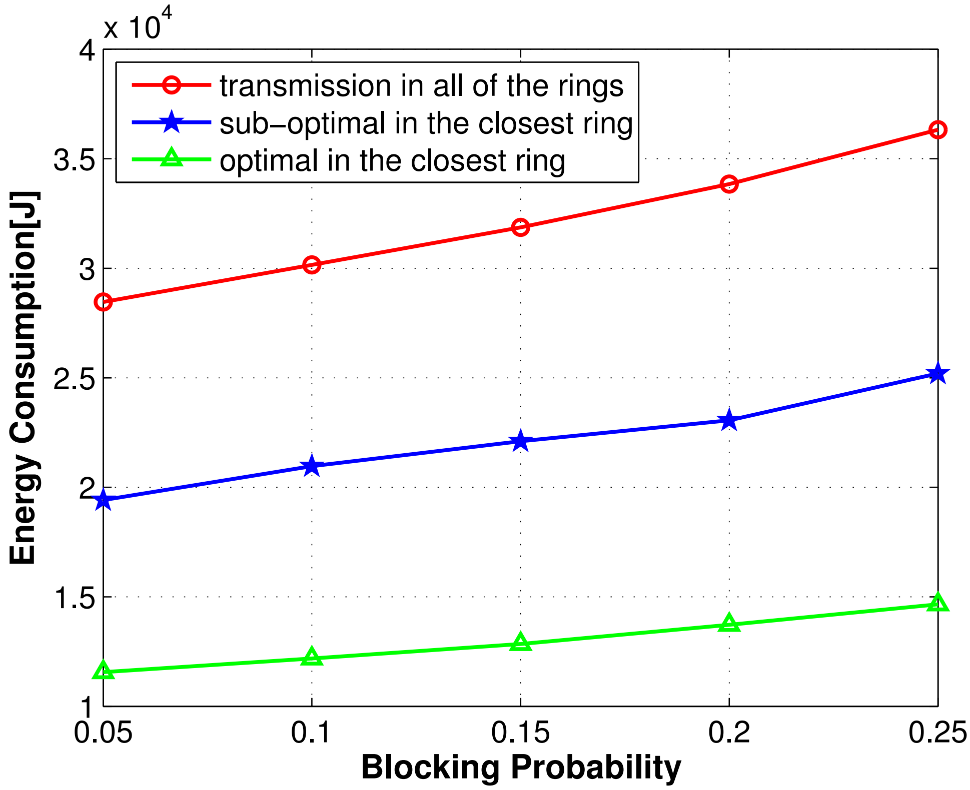


Figure 3.7: Energy consumption of transmission in different blocking probability thresholds for one SU

Concerning the energy consumption, Figure 3.7 presents three different cases of energy cost focusing on one SU to be served. Firstly, when the wireless node is moving towards the BS, we compute the energy consumption in all the six rings of the BS. The second case shown is the situation when the SU buffers the message to be transmitted to the BS and moves into the last ring for data transmission, which the SU occupies all of the bandwidth of the BS. Lastly, based on the second case, the proposed scheme is depicted where the SU has moved into the last ring of the BS, as the bandwidth increasing, the SU queuing system can accommodate an increased number of SU messages under a threshold of blocking traffic rate, which brings all the capabilities into full play. Here, we must emphasize that in the situation of the first and third case, the queuing system works under an identical blocking probability threshold to restrict the message number in the queue. As a result, intuitively speaking, the optimal one (the third case) has the potential of increasing the energy efficiency due to the fact that it can process a higher number of SU message simultaneously; thus spending less time to transmit the message comparing to the sub-optimal one (considering messages of the same size).

From Figure 3.7, it is shown that, as the threshold of blocking traffic probability is set to be 10%, the optimal schemes only dissipate approximately 2/5 energy (Joule) with equal amounts of data transmission. It is noteworthy that, the energy cost of optimal schemes fluctuates narrowly among these three schemes under different blocking probability thresholds. The reason is that the optimal schemes provide the highest data throughput and accommodate largest number of SU messages, which leads to the shortest time duration for wireless transmission and circuit cost.

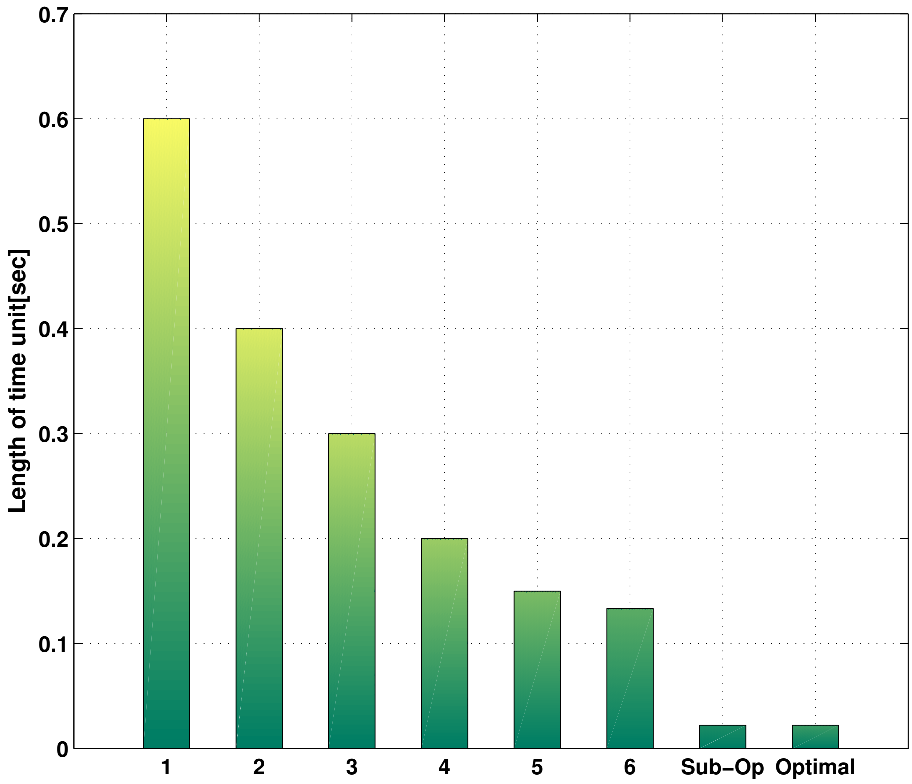


Figure 3.8: The length of time unit in different rings

It is assumed that the length of each message is identical. Due to the different throughput of wireless transmission in different rings of area, the length of time units that the wireless nodes could send one message to BS should be different as shown in Figure 3.8. Meanwhile, with the monotonically increasing values of blocking probability threshold, the number of SU message that can be accommodated in the queuing systems will rise in varying degrees. Compared to the first two cases, the optimal schemes can

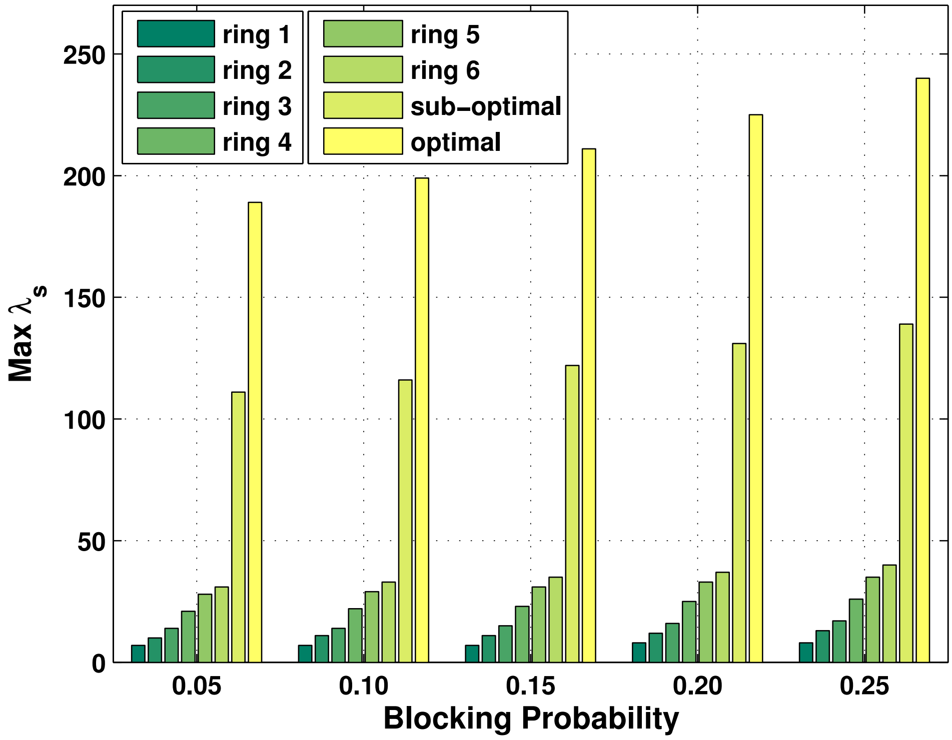


FIGURE 3.9: The max value of λ_s in different rings under different blocking probability thresholds

accommodate significantly more messages in the queuing systems simultaneously as shown in Figure 3.9.

Additionally, the scenario investigated above can be extended for the case where there are multiple SUs transmitting messages simultaneously. For illustration purposes, we set a blocking traffic threshold $p_{thres} = 0.10$ for the SU queues in this simulation as well. As presented in Figure 3.10, the trends of three curves are almost the same as before and the proposed scheme outperforms the other schemes.

3.5 Summary

In Chapter 3, in order to capitalize the delay tolerance of applications, a theoretical framework is developed for energy efficiency schemes that can maximize spectrum utilization and SU throughput. For cellular interface, the downlink/uplink rates of mobile devices can dynamically be adjusted according to the modulation and coding scheme

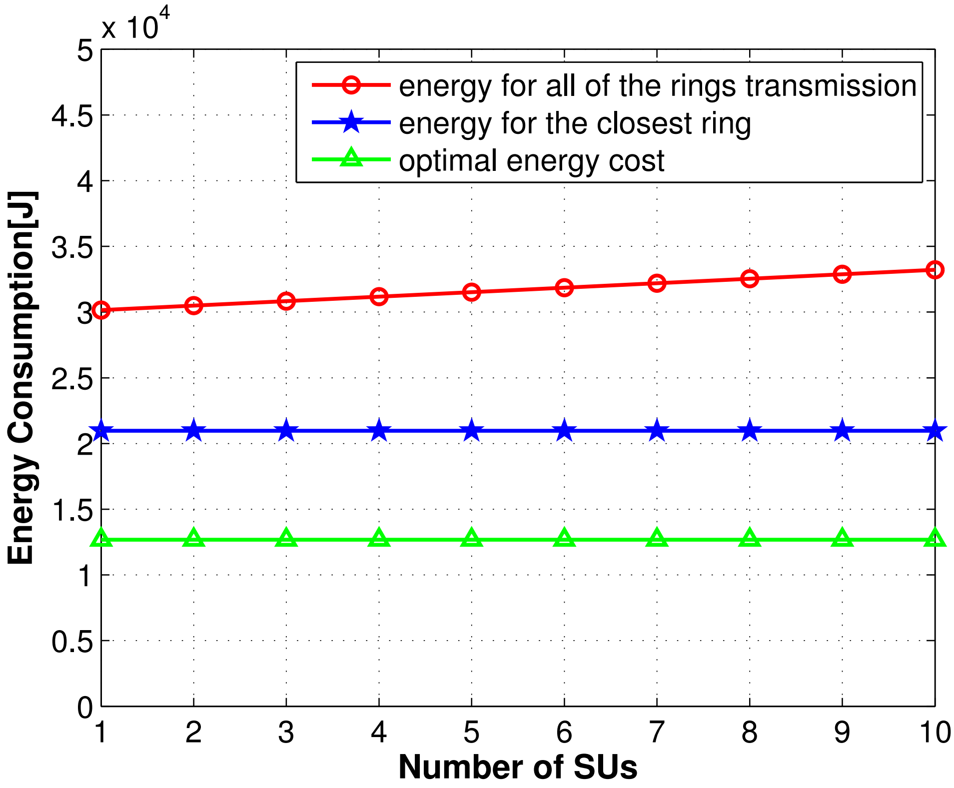


FIGURE 3.10: energy consumption of transmission with increasing SU number

based on the distance between BS and mobile user, which has been presented in subsection 3.1.1. Therefore when the mobile user is far away from the serving BS, it is necessary to look for alternatives to reduce the wireless transmission cost. Currently, as many portions of the TV spectrum are not in use for a significant period of time, it is implied that the existence of plenty of spectrum opportunities can be potentially exploited for wireless transmission. Hence, once the mobile devices are equipped with a dual or multi mode antenna that is connected to cellular networks and other networks such as Wi-Fi and White-Fi, the SUs could switch between these networks to seek and use any licensed spectrum bands as long as they do not cause interference to the PUs.

By considering the distribution of the SU traffic loads and PU connections that would be emerging stochastically, the CR system will contact a trusted database for historical information about PU traffic at a specific location and time duration so as to estimate the probability for the SU connections. With the aid of the estimated available probability concerning location and time duration, the SU traffic can access the primary channels more efficiently and increase the channel utilization. Based on the estimation of PU traffic within a specialized relatively long-term duration, and the analysis of probabilities for vacant channels and time slots, an M/M/K/L queuing system is devised in subsection 3.1.3 to estimate the SU traffic capability that the system can serve simultaneously. By considering an M/M/K/L queue where SU connections compete for K frequency channels in N concentric rings with different modulation and coding schemes within the BS coverage, the traffic of the SUs can be inputted to the queues for the SU connections. If the number of SUs is large, the input traffic of the virtual queues can be modeled as a Poisson process, where L is the finite number of waiting positions for each queue.

Finally, a scheme is proposed for delay-tolerant mobile applications in CR networks with a central aim of reducing the energy consumption. When non-real-time applications are buffered in the SU systems, the system is recommended to delay the data transmission to an area close to the BS, thereby maximizing the throughput potential whilst reducing the overall energy cost of wireless transmission under several constraints. The analysis and numerical results are presented in section 3.4, which reveal that the proposed optimal wireless transmission schemes can accommodate significantly more messages in the queue simultaneously under the same blocking threshold. Also, due to the highest data throughput of the proposed schemes, the overall energy cost of wireless transmission can also be significantly decreased. In addition, when the scenario is extended to multiple SUs cases, the proposed scheme still outperforms the other schemes.

Chapter 4

Diversity of User Mobility in Wireless Transmission Scheduling

In Chapter 4, the stochastic characteristics of user mobility is studied which should include direction changes and variable speed of the mobile users. As the time spent by a subscriber unit in the coverage of Wi-Fi/White-Fi hotspots strongly affects the overall efficiency of wireless transmission, a realistic movement model is analyzed which focuses on the velocity of mobile users, direction changes, and route selection distributions. Thereafter, the CRT (cell residence time) in wireless hotspots is utilized to optimize the schemes that can provide better performance in terms of energy saving, application requirements, and switching cost of different radios in embedded systems. Furthermore, a strategy that deploys roadside infrastructure Wi-Fi/White-Fi APs at a given location is introduced to assist the data delivery for mobile devices according to corresponding the on-line service features of mobile applications. In this framework, a probabilistic analysis is incorporated for the route diversity, and the CTR of user mobility. For the delay-tolerant mobile Internet applications, the schemes proposed in Chapter 4 can effectively extend battery lifetime by making selective use of the available high-rate roadside wireless hotspots.

4.1 Mobility Analysis

Since wireless node mobility greatly affects the performance of mobile wireless networks, a realistic mobility model is critically required to provide a deep analysis on the impact of the terminal mobility, which includes user velocity, direction changes distribution, and route selection distribution. The spatial and temporal characteristics of RWP (Random Waypoint) are analyzed as a discrete-time stochastic process in [122]. A time-variant mobility model is proposed in [123] to characterize mobility pattern based on data of daily activities. It is assumed that when a mobile device is engaged in a wireless communication context, the system of mobile device system must be able to periodically update its current location to the heterogeneous network. Hence the system of heterogeneous network can determine its served mobile users current access point in order to pre-computed and predict their route properly for incoming wireless transmission. Location updating procedure is that the mobile users will periodically inform the system about changes of their current status including speed changes, direction changes and pre-determined path changes.

Historical information concerning the movement of vehicles can be used to predict the location that a vehicle is more likely to be found or the action that a vehicle is more likely to take. There are some subsidiary methods that can be used to aid realistic route and location estimation of mobile users, such as path selection probabilities, and time-dependent mobility models of mobile users (both vehicles and pedestrians), which would greatly facilitate the design of cost-efficient wireless transmission strategies that meet demands of mobile applications. Furthermore, some road rules which also can be utilized to aid route prediction and location estimation.

1) The speed limit of a given road can be used to compute the cell residence time in one Micro-cell, the distance between mobile users and the serving BS, and the uplink/downlink rate via multiple wireless accesses. In the meantime, other related factors, such as traffic congestion, traffic light, and performance limitations of different vehicles (car, bus or truck), could affect the likelihood of a hypothesized trajectory traveling along the given road.

2) There are several restrictions when mobile users are traveling at a certain street to be along with, such as one-way restrictions, no left turns, no U turns, vehicle type

limitations (height and weight), pedestrian-only street, give-way sign at an intersection that mobile users must slow down, and bridge/road weight limit to bear, which could be important factors for wireless transmission strategies.

3) Finally, many empirical information can be collected to aid the prediction of movement for mobile users. A considerable portion of vehicular/pedestrian journeys are repeatable. For example, the mobile users on the public transportation (train, ship, and bus) have fixed routes. In addition, the route and time for every week, among their home, school, work place, super market, shopping center, are trackable and stable with some rules to be followed. These useful information regarding daily route of vehicles and pedestrians could be updated regularly. Since the historical information could be possible to more accurately locate a mobile user, once this priori information are available for the prediction of mobile transmission, there would be a significant energy saving for wireless transmission.

4.2 Macro and Micro Cells

LTE Heterogeneous network offers a significant capacity improvement in cellular networks by reusing frequency resources within a Macro-cell coverage area, where a couple of Femto nodes are deployed for higher capacity [79]. The Macro-cell provides basic coverage in Heterogeneous networks, while the Femto nodes are placed randomly inside a Macro-cell to increase wireless transmission capacity and prevent coverage holes [80]. Therefore, a future capacity solution for LTE is envisaged by utilizing small cells (Phantom Cells) to carry user traffic and offer good mobility support [81], where a new interface (X3) is required to perform the master-slave relationship between the Macro Node and the Phantom Nodes. As the phantom Cells are deployed in a different carrier frequency with the Macro-cell, the UEs (User Equipment) need to transmit discovery signals which will be synchronized with the Macro-cell in a periodic manner for fast and power-efficient detection. The research in [81] focuses on L1/L2 mobility procedure in a continuous Phantom cell layout, which the handovers are executed between Micro-cells (Phantom cell).

4.2.1 Vertical Handover

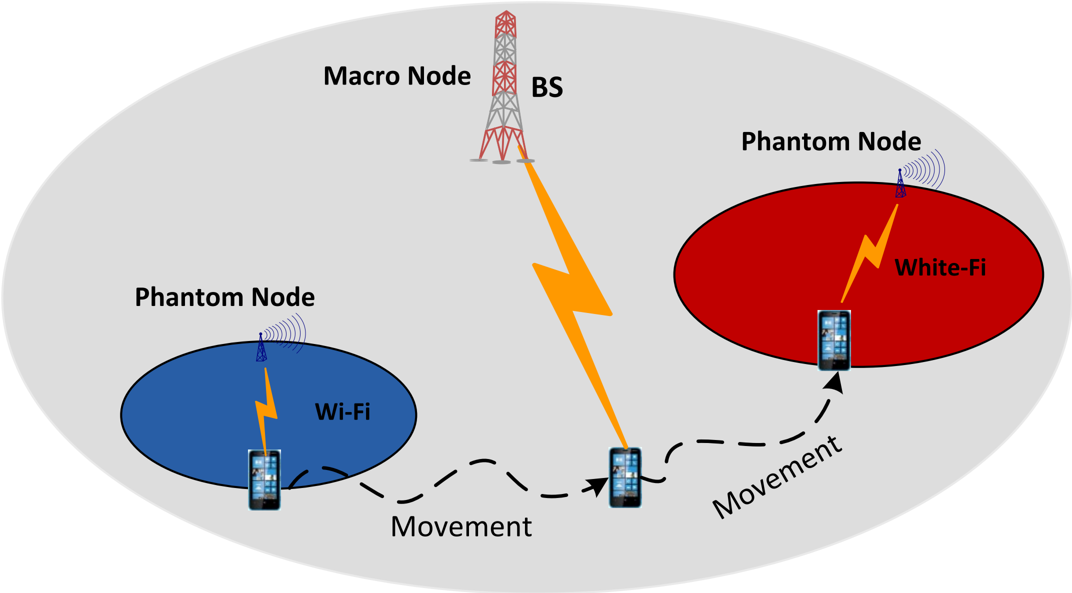


FIGURE 4.1: Vertical handover in heterogeneous networks

Vertical handoff process allows the mobility of MNs (Mobile Nodes) among access points supporting different network technologies as shown in Figure 4.1. Handovers from a serving network to a target network should be fast so that mobile users can continue receiving their services seamlessly. Supporting voice and interactive multimedia with continuous mobility implies that the handover latency should not exceed 50 ms to prevent excessive jitter [99]. As stated in [86], a handover delay of less than 200 ms is acceptable for supporting real-time services. A test-bed implementation based on a RSSI (Received Signal Strength Indicator) and hybrid RSSI/goodput VHO algorithm is proposed in [89], which also experiences a long handover latency between commercial WiFi and UMTS.

To satisfy the needs of mobile users, MIPv6 (Mobile IPv6) protocol has been proposed by the Mobile IP IETF (Internet Engineering Task Force) to solve the issue of IP mobility, which allows a MN to be identified by a single IP address even though the MN may move its physical point of attachment from one network to another [124]. Various messages are exchanged between MNs and a home agent which suffers from inefficient utilization of resources such as long handover latency, high packet loss and signaling overhead. A testbed is implemented in [91], which the MN has three access interfaces (WLAN, GPRS and TD-SCDMA) supporting MIPv6. And a wireless heterogeneous network is established to evaluate the performance of vertical handoff delay.

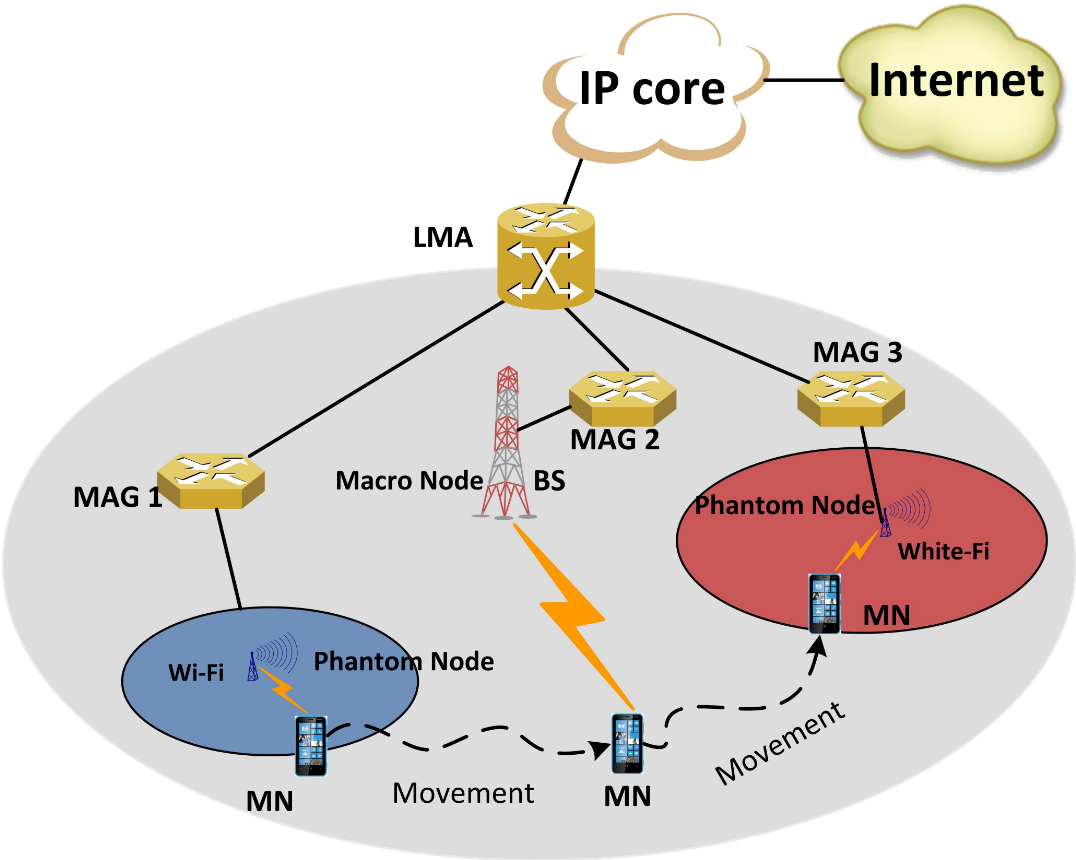


Figure 4.2: PMIPv6 architecture in heterogeneous wireless networks

PMIPv6 (Proxy Mobile IPv6) protocol is an enhanced version to MIPv6, which supports IP mobility operation without requiring the participation of MNs in mobility task [96]. Hence, It is able to reduce handover latency. PMIPv6 introduces a special access router, called MAG (Mobile Access Gateway), which ensures an MN to continue the same IP address configuration as long as it roams within the same PMIPv6 domain [98]. An LMA (Local Mobility Anchor) is in the backbone network maintaining a collection of routes for MNs within the located mobility management domain as shown in Figure 4.2. Packets for an individual MN are routed to and from the MN through tunnels between the LMA and the MAG. When a MN moves from one link to another upon handover, the MAG will send a route update to the LMA. A handover mechanism is proposed in [98] to improve the performance of PMIPv6 by incorporating an HC (Handover Coordinator) to enhance handover performance and reduce handover delay. The work in [90] evaluates the benefits of PMIPv6 protocol regarding handover performance between WLAN and HSDPA in a real test-bed. During an efficient handover the IP address should not be changed. An experimental evaluation of PMIPv6 is presented in [93], which analyzes the

TCP/UDP performance and the handover latency of PMIPv6 from WLAN to WLAN.

Since the variety of vertical handoff decision algorithms is too complex for the immediate use of smartphones, an embedded VHDA (Vertical Handoff Decision Algorithm) is proposed in [87] for seamless data transmission upon vertical handoff, which can be readily used on latest smartphones without any changes in the network infrastructure. Behaviors of 3G/WLAN seamless handover on mobile applications (FTP/HTTP downloads, video/audio streaming, Skype) are evaluated based on a Nokia N900 test-bed in [125], which utilizes a vastly simplified PMIPv6-like handover procedure. From Table 4.1, it is shown that the handover latency is extremely long in their test-bed experiment.

TABLE 4.1: List of Handover Simulations

Protocols	**Platform**	**Interfaces**	**Latency**
MIPv6	test-bed	WLAN, GPRS, TD-SCDMA	WLAN to TD-SCDMA: 0.91s [91]
RSSI/goodput	test-bed	WiFi, UMTS	WiFi to UMTS: 4.13s UMTS to WiFi: 5.43s [89]
PMIPv6, FMIPv6, PMIPv6-HC	NS-2	WiFi, WiMAX	WiFi to WiMAX: 0.58s (PMIPv6) [98]
PMIPv6, MIPv6	test-bed	WLAN, HSDPA	N/A [90]
PMIPv6	test-bed	WLAN to WLAN	higher than 100ms for TCP [93]
VHDA	test-bed	WiFi, UMTS/HSPA	UMTS to WiFi: 140ms WiFi to UMTS: 230ms [87]
Simplified PMIPv6	test-bed	3G, WLAN	3G to WLAN: 11.72s [125]

4.2.2 Cell Residence Time

The cell residence time (CRT) in one hotspot area (such as Wi-Fi, White-Fi) is the time interval from the start point of the wireless transmission to the time that the mobile user moves out of the cell. The research in [102] examines the sensitiveness of mobile network performance in a discrete event simulation, which the handoff dwell time is following an exponential and truncated Gaussian distribution. The work in [106] evaluates the parameters of handover margin and averaging window by approximating the handover cell residence time as a generalized gamma distribution. Similarly, mathematical formulations are developed for the systematic tracking of the random mobility models in [100], which cell resistance time is approximated by the generalized gamma distribution.

The effect of mobile cellular network performance has been quantified in different distribution cases that cell residence time is Exponential, Erlang, Gamma, Uniform, Weibull and Deterministic [101].

When the mobility of mobile users can be well predicted, systems can have a prior knowledge of the mobile users' destination. In realistic cellular mobile networks though, the time probability distribution takes into account the fact that some users may choose to take a longer route to the destination. When a mobile user is moving at an intersection, the destination, road structure and traffic conditions will influence the direction selection of the user. Moreover, a mobile user will move along different paths with different speeds, depending on the road network pattern and rules, which leads to the different time spent on these roads. Consequently, the mobility modeling should include the direction changes at a crossroad and the speed of the mobile users. To this end, we assume that the cell residence time follows a negative exponential distribution.

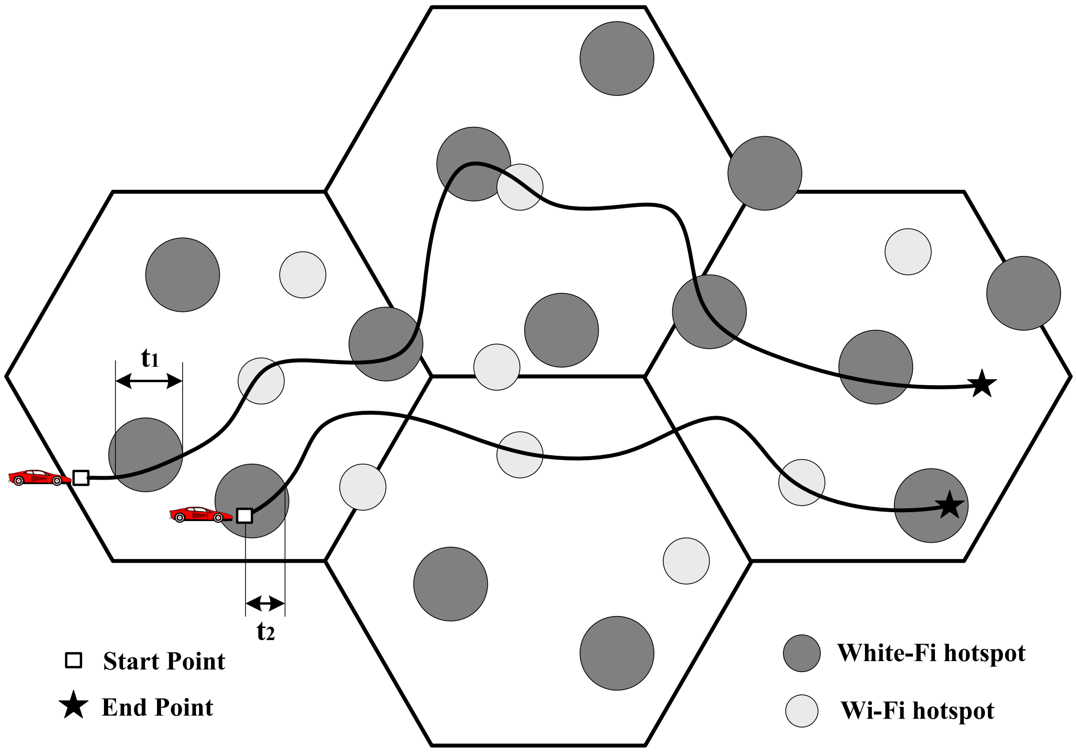


FIGURE 4.3: Graphical representation of node mobility

The time spent and distance traveled by a mobile user in one Micro-cell depends on its average speed, changes of direction, route selection, route shape and traffic situations. In urban roads environment, most mobile users (vehicles, pedestrians) tend to have a relatively constant speed and travel along relatively fixed trajectories. Depending on whether one file transmission is completed within the coverage area of one access cell (Wi-Fi or White-Fi) or handed over to cellular networks (LTE), two different cell residence times can be specified: (1) the remaining cell residence time, and (2) the handover cell residence time, noted as t_2 and t_1 respectively and shown in Figure 4.3. If the start point or the end point of a mobile user within one hotspot, the sojourn time in this hotspot should be the case of remaining cell residence time $E[T_n]$; the case of handover cell residence time $E[T_h]$ is the time duration to move in and out from the cell. These two cases of expected mean residence time for an arbitrary mobile user in a cell have been theoretically evaluated in [126] and they can be described by the following equations,

$$E[T_n] = \frac{8R}{3\pi E[v]} \tag{4.1}$$

$$E[T_h] = \frac{\pi R}{2E[v]} \tag{4.2}$$

Where R denotes the radius of the cell and v is the average speed of the mobile user. Moreover, the exponential distribution is used extensively to characterize the CRT because of its analytical tractability. Since the residence time of mobile users could be different according to their trajectory and hotspot coverage, the systems should ensure that there would be enough probabilities in the wireless hotspot coverage to assure for a complete file delivery.

4.2.3 Wireless Transmission Duration

The wireless transmission duration times are assumed to be independent and identically distributed random variables, which strongly depend on the location and mobility of mobile users.

TABLE 4.2: Parameters for Wi-Fi/White-Fi Hotspots

Hotspots	**Distribution**	**Mean $E[T_h]$ (sec)**
$Wi\text{-}Fi_1$	Gaussian Distribution	10
$White\text{-}Fi_1$	Gaussian Distribution	27
$White\text{-}Fi_2$	Gaussian Distribution	22
$Wi\text{-}Fi_2$	Gaussian Distribution	15

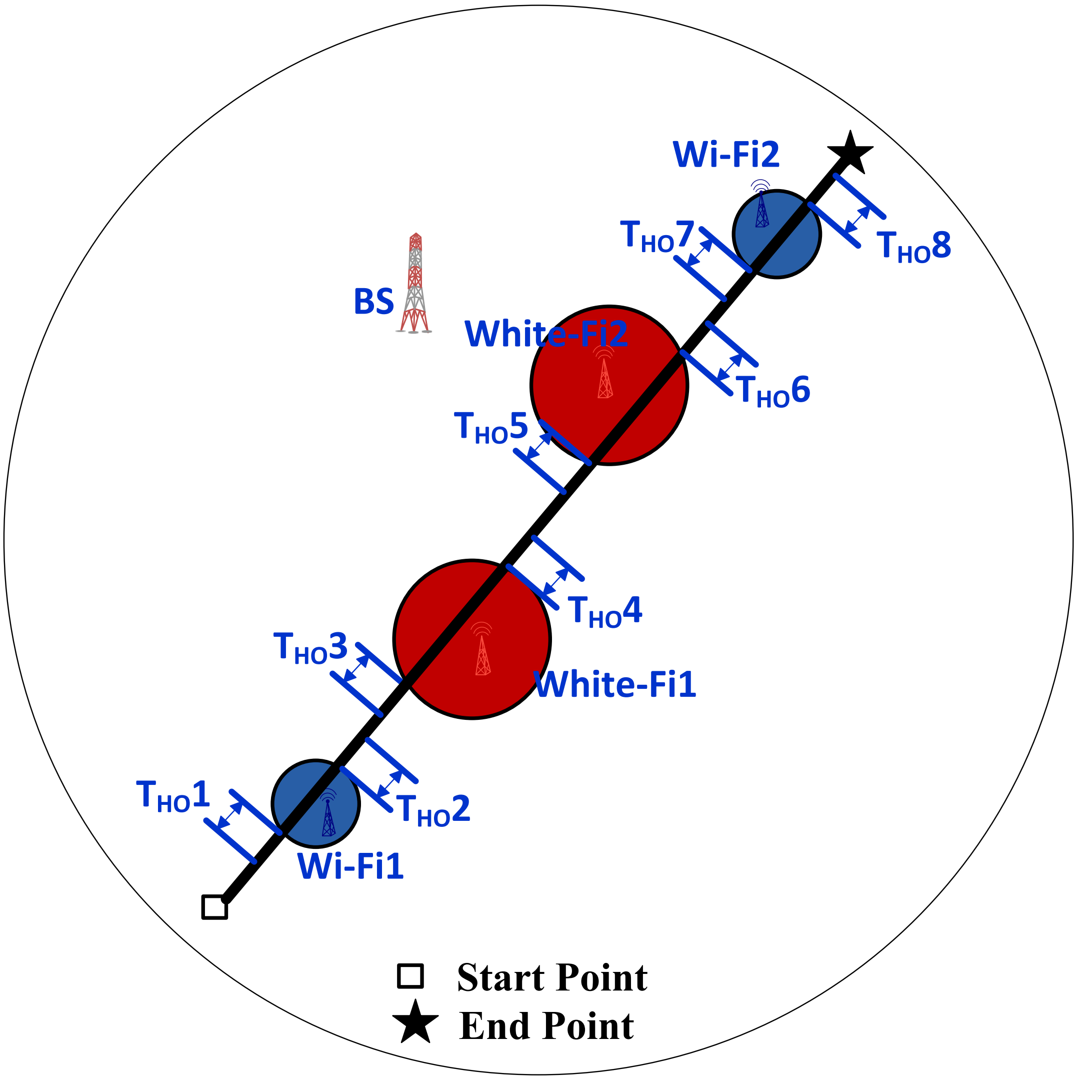


FIGURE 4.4: A typical route of mobile users

A arbitrary vehicle trajectory of a mobile user within the coverage of cellular networks and hotspots is shown in Figure 4.4 under the assumption that the wireless transmission will terminate after K hotspots. Let $T_{HOi}(i = 1, 2, \cdots, 2K)$ be the handover delay when the wireless access of the mobile user switches from one to another. $T_{CRj}(j = 1, 2, \cdots, K)$ is the cell residence time under the high-speed hotspot coverage and T_{Sum} is the time interval between the start point and the end point of the mobile user. Let $T_{CellRes}$ be the time duration that the mobile user can access the cellular networks without the coverage of Wi-Fi/White-Fi hotspots.

$$T_{Sum} = \sum_{i=1}^{2K} T_{HOi} + \sum_{j=1}^{K} T_{CRj} + T_{CellRes} \tag{4.3}$$

Suppose that random variables (RVS) T_{HOi} for any i is exponentially distributed with mean $1/\eta$. And the cell residence time T_{CRj}, which is the time duration that a user stays in the coverage of a hotspot cell, follows Gaussian distribution with mean $E[T_h]$ as presented in Table 4.2. Consider a heterogeneous wireless network including one BS, two Wi-Fi hotspots and two White-Fi hotspots as shown in Figure 4.4, we assume the radius of the macro cell (BS) to be 1000 meters, and the radius of the White-Fi hotspots and Wi-Fi hotspots to be 100 and 50 meters respectively. From the start point of this map, there are ten vehicles moving towards the end point of this map along with the same route as presented in Figure 4.4. In this scenario, each vehicle is moving continuously for 250 seconds. One case is the mobile users on these vehicles only transmit data via cellular networks. Second case is that the mobile nodes switch between cellular networks and Wi-Fi/White-Fi interfaces, which will generate 8 times handovers and access Wi-Fi/White-Fi hotspots for high-rate wireless transmission. Figure 4.5 illustrates the simulation results and compares the above two cases, and it is shown that during the 250 seconds, the data sizes downloaded by 100 mobile nodes are all around 466 Mbits by only cellular access. On the other hand, the data sizes downloaded from second case vary considerably because of the variation of cell residence time in these 4 Wi-Fi/White-Fi hotspots. From Figure 4.5, it is observed that, if the mobile nodes always execute the handovers between Micro-cells and Macro-cells, the data size downloaded by mobile nodes could be significantly increased.

4.3 Long Term Energy Saving

In this scenario, it is assumed that the road segment a vehicle is moving along is covered by a cellular BS with two White-Fi hotspots and two Wi-Fi hotspots in this domain as shown in Figure 4.6. Let $\Gamma = \{\tau_i, \tau_{i+1}, \cdots, \tau_j\}$ denote the decomposed time slots in which mobile users is moving along a road. Let $Q(t) = \{q_i, q_{i+1}, \cdots, q_j\}$ represent the energy consumed from embedded peripherals, such as CPU, Graphics, backlight and

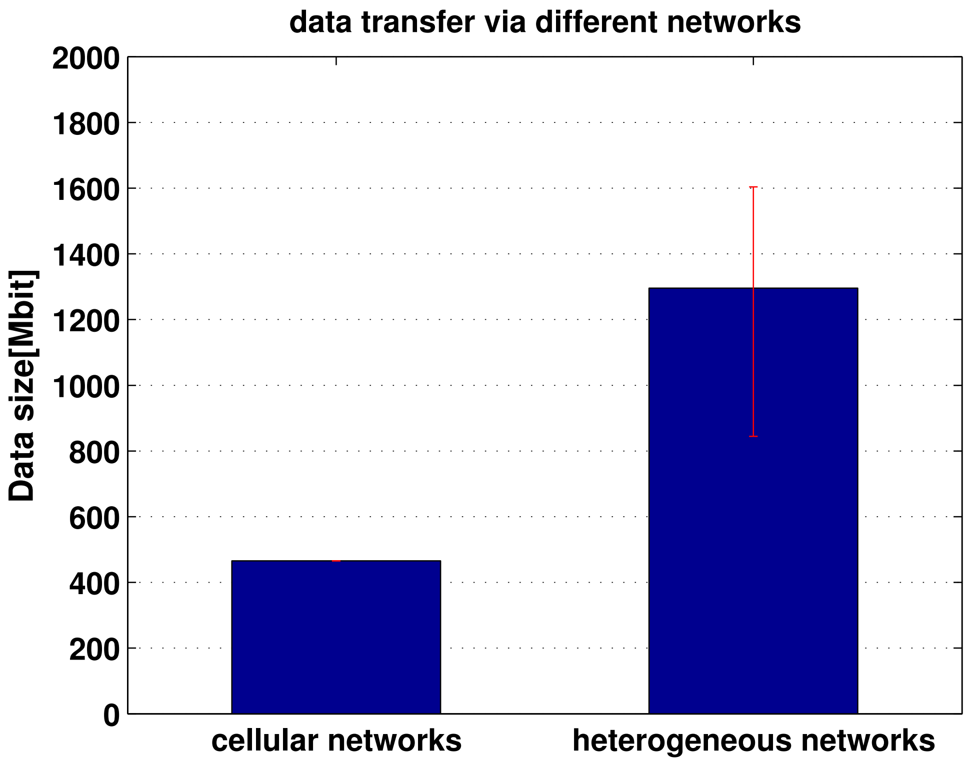


FIGURE 4.5: The different sizes of wireless transmission via two different networks (10 sample vehicles)

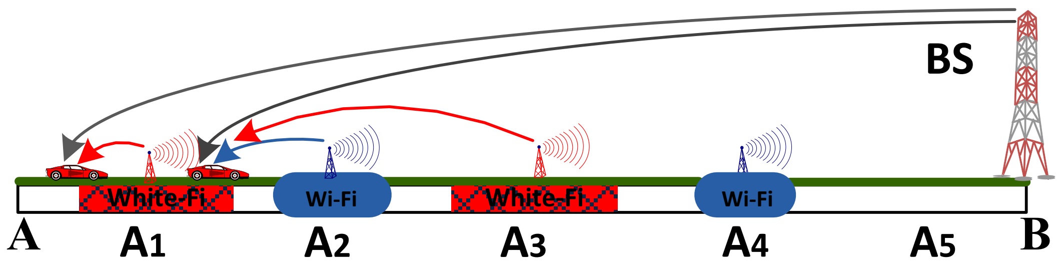


FIGURE 4.6: A general model used for evaluating energy saving of smartphones

storage devices. Our goal is to find the minimized energy cost for a fixed file size transfer over wireless access as follows,

$$E_{min} = \min_{s_t \in S} \sum_{t=\tau_i}^{\tau_j} \left\{ E_{cell}(t) + E_{TV}(t) + E_W(t) + Q(t) \right\} \tag{4.4}$$

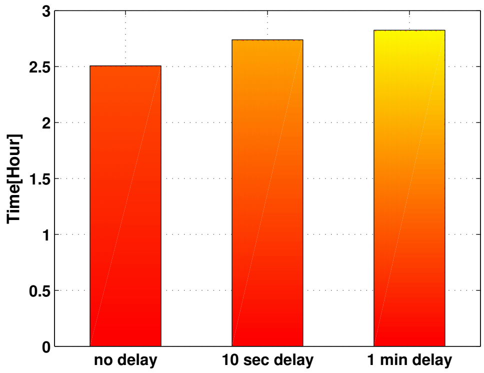
Where s_t represents the strategy at time slot t, and S is the set of possible strategies for wireless transmission. E_{TV} and E_W represent the energy consumption for White-Fi and Wi-Fi components of mobile devices respectively. E_{min} is the overall energy cost for the embedded system in the mobile user end. Our proposed strategies strive to achieve energy efficiency for mobile users across a long-term average, thereby prolonging the battery lifetime. Table 4.3 summarizes the energy cost of main peripherals in a standard embedded system of mobile devices.

TABLE 4.3: Power Requirements for Peripherals in Mobile Devices

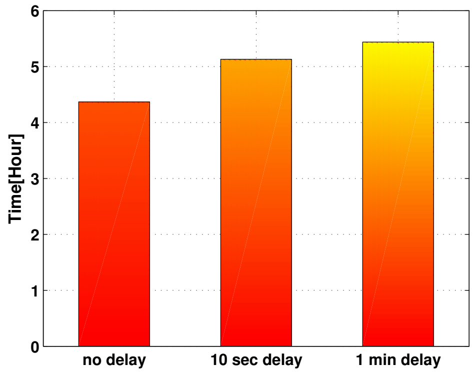
Component	**Action**	**Value**
CPU Usage	100%	$612mW$
	50%	$462mW$ [5]
	2%	$55mW$
Backlight	White background 100%	$527.05mW$ [5]
	Screensaver mode	$13.86mW$
Graphics	Video playback	$110mW$ [127]
	other	$80mW$
DRAM	Average Energy(nJ)/read	$56nJ(/64bytes)$
	Average Energy(nJ)/write	$61nJ(/64bytes)$ [128]
	DRAM Precharge Power	$560.0mW$ [5]
Wi-Fi	Downloading	$1450mW$ [5]
	In disconnection	$135mW$
3G	Downloading	$1400mW$ [5]
	Idle	$58mW$

In this case study, the effect of this scheme on long-term usage will be considered. Previous statistics have shown that the average multimedia video size is $10MB$ on YouTube, which is a nominal value for such video files [129]. Hence we assume that the average length of a YouTube video is 4 minutes and 12 seconds, so in one cell, the mobile user will watch 250 seconds video which translates to about $100Mb$ downloading from wireless radios on average.

Additionally, it is assumed that the mobile devices have a lithium-ion battery with a capacity of 1200mAh, 3.7V. Figure 4.7 shows a comparison of the battery lifetime of two mobile applications (video playback and audio playback) in three different scenarios. In the first scenario, when the mobile users make a request for the application, the system will start wireless transmission immediately without any delay. In the second and last scenarios, the mobile systems would delay the wireless transmission 10 seconds and 1 minute respectively. In the case of video playback, as seen in Figure 4.7A, the strategies that allow delay for mobile application can achieve drastic energy savings. As a result,



(A) Video Playback



(B) Audio Playback

FIGURE 4.7: Battery life for different mobile applications

the 10 sec delay constraint strategy and 1 minute delay constraint strategy could extend the battery lifetime by 9.3% and 12.7% respectively.

Similarly, in the audio playback case, it is assumed that 100Mb data size will be downloaded via wireless access every 250 seconds. As can be seen in Figure 4.7B, the proposed schemes which allow delay on the transmission perform better than the non-delay schemes due to the fact that in the audio playback scenario, the embedded system will have less CPU utilization with the backlight off, which takes a significant portion of overall energy consumption in the video playback scenario. Consequently, in this case, the battery lifetime will be extended up to 24.4% compared to the non-delay strategies.

4.4 Summary

Since battery technology does not well match with the development of Internet applications, the energy limitations have been a direct impact on the time that mobile devices are operational. As previous investigations disclose that data transmission via wireless radios is a dominant cost in mobile devices, it is important to intelligently make selective use of the high-rate wireless accesses with stochastic features. Since the time spent by a subscriber unit in the coverage area of a Micro-cell would significantly affect the overall efficiency of wireless transmission, realistic movement of mobile users is analyzed in section 4.1 including velocity, direction changes, and route selection distribution.

Vertical handover process allows the mobile terminals switch between high-coverage Macro-cells and high-rate Micro-cells for cost-efficient wireless transmission. To fulfill the requirements of mobile applications, a framework of heterogeneous networks has been introduced to provide the seamless wireless services for mobile users, which the cell residence time in Micro-cells are of major importance with respect to service quality evaluation. According to the review presented in subsection 4.2.1, there is a handover cost such as handover delay when the mobile devices switch between cellular networks and White-Fi/Wi-Fi hotspots. According to the fact that wireless transmission could be initiated at any point within the Macro- or Micro- cells along the route of the vehicle/pedestrian, two different cases including the remaining or the handover residence time has been discussed in subsection 4.2.2. Based on an arbitrary vehicle trajectory in a typical city, two strategies of wireless transmission have been compared in subsection 4.2.3, although there is handover cost introduced by switching between Micro-cells and Macro-cells, the data size downloaded by mobile devices could be significantly increased by utilizing the high-speed hotspots if the handover cost is small enough.

In the meantime, as the data transmission via wireless radios is a dominant energy consumption in mobile devices, in order to avoid the drain of mobile device batteries, a novel strategy is proposed in section 4.3 for battery lifetime extension that capitalizes on the delay tolerance of mobile Internet applications by making selective use of the available high-rate roadside wireless hotspots (Wi-Fi/White-Fi). Experiments based on theoretic and realistic analysis reveal that impressive energy savings can be achieved, which the battery lifetime could be extended up to approximately 25% for the applications with a high delay tolerance. In the future, the mobility model that distinguishes pedestrians and vehicles could better characterize the movements of mobile users to deeply investigate the energy saving for these two cases, and the effects of changes in direction and speed must be considered together.

Chapter 5

Energy Savings in Mobile Embedded Systems

Significant developments in streaming multimedia applications has been witnessed recently for digital devices, such as smartphones, Tablet PC, and PDAs; these efforts have remarkably re-shaped user experience for online wireless multimedia. Mobile users expect that the convenience of online services, such as news, sports, or entertainment clips, can be experienced anytime and anywhere by simply clicking on a button of digital devices. The increased usage of smartphones and the rich ecosystem of Internet applications are having a severe effect on the recharging cycles of devices due to the increased levels of energy consumption and limitations of battery technology; eventually battery capacity becomes a key constraint due to the size and weight of digital devices. To this end, it is critical to develop innovative strategies in order to manage the wireless embedded systems efficiently, thereby prolonging the battery lifetime and enhancing overall user experience.

First of all, Chapter 5 will focus on the overall energy saving in mobile embedded systems, which will include the analysis of key embedded peripherals such as CPU, Graphics, backlight and storage devices (DARM, FLASH). Under the inherent stochasticity of available of transmission opportunities, the challenge is to select an optimized time duration to launch the data transmission in order to minimize the overall energy cost while satisfying the information delay constraints (with focus on delay-tolerant messages). To this end, a novel relaying scheme to deal with the problem of energy-delay trade-off, based on optimal stopping programming (OSP), is presented in section 5.3, whilst exploring different candidate solutions for data transmission. Apart from that, recent studies reveal that users may want to access the same popular video content multiple times which can be energy inefficient if it is always streamed to the users. Hence in section 5.4 different policies are studied in terms of whether or not content should be stored in the device by taking into account the probability of re-using the content, the energy consumption of DRAM (if it is to be stored) and the energy consumption to stream the content again.

5.1 Delay Cost

We assume that a vehicle node has a file of F Mbits in the buffer (DRAM) for transmission at the first available time slot. In a similar manner to the assumptions made in [33], all the data has to be transmitted to the BS before a hard deadline (T). Intuitively speaking, simply increasing the average delay time of packets can effectively reduce the energy consumption of the model. Suppose that a chunk of the information file is segmented into several packets of length as $\{l_0, l_1, \cdots, l_m\}$. Let $\Gamma = \{\tau_0, \tau_1, \tau_2, \cdots, \tau_j\}$ denote the time slots along the whole route of vehicle. So the aim of our model is to choose the best time duration $\{\tau_k, \tau_{k+1}, \cdots, \tau_{k+m}\}$ from all the time slots above for transmitting the data to the BS in order to minimize the overall energy consumption, where τ_k is the time slot to launch the message transmission, and τ_{k+m} is the time slot that complete the transmission. At each time slot a vehicle node can forward some of them, e.g., l_1, to the BS. The average time delay for all the data packets is given by,

$$T_{delay} = \sum_{i=0}^{k-1} \tau_i + \frac{1}{F} \sum_{i=0}^{m} \tau_{k+i} \cdot l_i \leq T \tag{5.1}$$

As seen in (Equation 5.1), we include a hard deadline T to restrict transmission time duration in the proposed transmission scheduling strategy.

5.2 Storage Cost

Broadly speaking the general architecture of a mobile device (such as smart phones, tablet computers, digital cameras, etc.) can be decomposed into the processing unit (CPU), the local Dynamic Random-Access Memory (DRAM) and flash/hard disk (HD-D). When a wireless terminal prepares to transmit data to other nodes, the data has to be ready in the local DRAM. However, if the system decides to wait for a short period to transmit, the data stored in DRAM would be transferred to internal storage devices (e.g., internal NAND flash) or external storage units (e.g., SD card, HDD) [130], which depends on the delay constraints and transmission strategies of mobile applications.

The research in [128] splits the DRAM power into two parts, namely operation power and background power. Apart from the command operating power, such as read/write, register power, and termination power, the background power is the power that depends solely on the power-down state and operating frequency. The conventional power management technique of DRAM mainly focuses on the trade-off between power consumption and wake-up latency. There are three major power states with decreasing power consumption: standby, power-down, and self-refresh. In [128], the DRAM energy required for each operation at two operational frequencies, 1333MHz and 800MHz, have been presented. And we include the energy cost under 1333MHz operational frequency in a typical server platform as a reference to show the differences of energy consumed by different operational modes in Table 5.1. In our scenario, memory designed for mobile platforms, such as LPDDR2, will be utilized to design the proposed energy efficiency schemes, which far more energy-efficient in contrast to PC/server DRAM [131][132].

Table 5.1: Storage Power Requirements in Different Operational Modes

Symbol	Metric	Value
P_{Nidle}	NAND flash Idle Power	$0.4mW$ [127]
e_{Nread}	NAND flash Read efficiency	$1.92nJ/bit$ [127]
e_{Nwrite}	NAND flash Write efficiency	$12.5nJ/bit$ [127]
$P_{HDDstdb}$	HDD Power in Standby Mode	$1.43W$ [133]
$P_{HDDidle}$	HDD Power in Idle Mode	$8.83W$ [133]
P_{HDDio}	HDD Read/Write Power	$15.63W$ [133]
e_{DRAMr}	DRAM Average energy(nJ)/read	$56nJ(/64bytes)$ [128]
e_{DRAMw}	DRAM Average energy(nJ)/write	$61nJ(/64bytes)$ [128]
P_{DRAMsf}	DRAM SelfRefresh Power	$0.92W$ [128]
P_{DRAMpd}	DRAM Precharge Fast Powerdown	$2.79W$ [128]
$P_{DRAMstdb}$	DRAM Precharge Standby	$4.66W$ [128]

Meanwhile, Flash memory has risen to prominence as an important component in mobile devices with large storage requirements, which supports three primary operations: erase, program and read [134] [135]. We use the known results from [127] to determine power consumption of NAND flash memories. Besides, as another important storage unit equipped in mobile computing devices (portable multimedia players and laptops), the energy consumption of the HDD should be aware for the system performance [133] [136]. As a general rule, the HDD has three operating states: standby, idle, and read/write. Table 5.1 presents the typical energy efficiency parameters of storage devices. Detailed descriptions of storage device operations can be found in datasheets, technical notes and papers [137] [138] [132] [139].

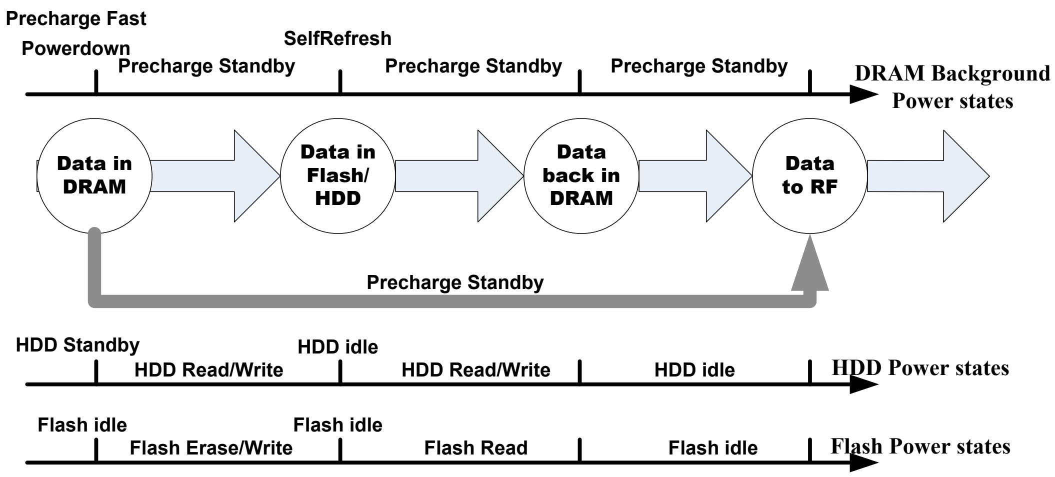


Figure 5.1: Power state machine of DRAM, NAND flash and HDD

In the numerical investigations the assumption is that the DRAM has three background power operating modes, namely the Self Refresh, Precharge Fast Powerdown, and Precharge Standby. Figure 5.1 shows the DRAM power states in different states for wireless transmission.

If as a result of the optimal policy the decision is not to transfer the message from the DRAM to NAND flash memory, the DRAM would enter into the Precharge Fast power-down state, in which the DRAM DLL is on, with relatively shorter exit latency in several power-down states. This is because DRAM would wake up at any time to transfer the stored data to the RF module. Once the embedded system starts transferring the data, the DRAM will repetitively execute operational commands (read/write array, read/write I/O, termination, activate and pre-charge) with the background power

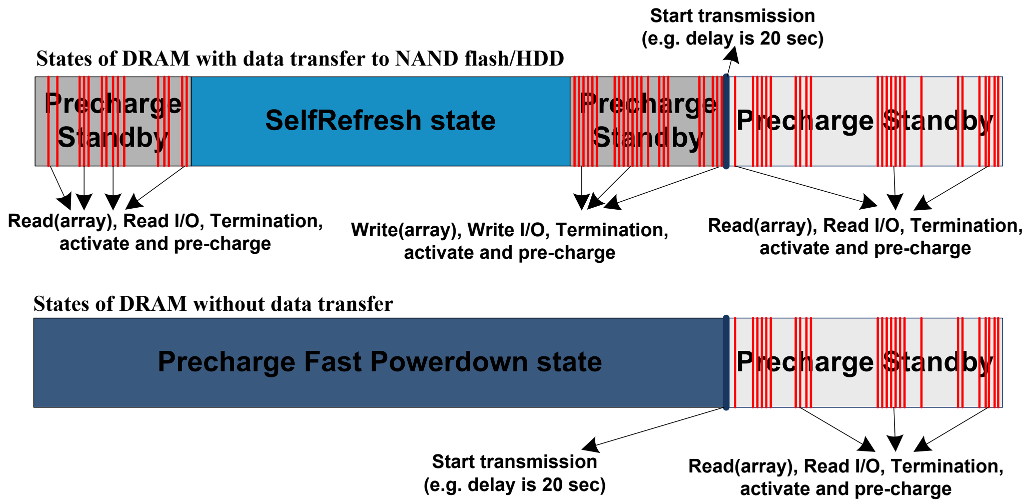


FIGURE 5.2: Operation and Background power state of DRAM

(Precharge Standby state). On the other hand, in the case that the optimal policy is to transfer the data from the DRAM to NAND flash, as illustrated in Figure 5.2, at first, the data would be transferred from the DRAM to NAND flash/HDD. During the transfer process, except for the operation commands, the DRAM would be in the Precharge Standby state for the sake of short wake-up latency, which will consume most power. After the completion of the transfer, the DRAM will enter into a self-refresh state, consuming least power with significantly higher exit latencies. Once the data return to the DRAM for data transfer to the RF module, the DRAM would enter into the Precharge Standby state again. In the mean time, the NAND flash will remain in the idle state except the duration of data transfer between the DRAM and NAND flash.

5.3 Optimal Stopping Problem

Based on the discussion above the objective of the energy efficiency optimization for the wireless transmission is to minimize the following expression.

$$\min_{\pi_t \in \{0,1\}, s_t \in S} \left\{ \sum_{t=\tau_0}^{\tau_{k-1}} Q(t) + \sum_{t=\tau_k}^{\tau_j} \omega_t(h_t) \cdot \pi_t \right\} \tag{5.2}$$

Let $\Gamma = \{\tau_0, \tau_1, \tau_2, \cdots, \tau_j\}$ denote the time slots in which the mobility of the terminal can be decomposed. The scheduling action is taken in time slot t: $\pi_t = [\pi_{\tau_0}, \pi_{\tau_1}, \cdots, \pi_{\tau_j}] \in$

$\{0,1\}$, π_t=1 if the data packet has been transmitted; we assume that the system would start the wireless transmission to the BS at time slot τ_k, thus $Q(t) = \{q_{\tau_0}, q_{\tau_1}, \cdots, q_{\tau_{k-1}}\}$ represents the energy consumed for buffering the data flow, such as energy cost of storage devices, and time delay. ω_t describes the transmission cost (e.g. RF power) at time slot t. In each time slot t, the user experiences a channel condition $h_t \in H$. $s_t \in S$ represents the vehicle's transmission strategy at time slot t, where S is the set of possible transmission strategies including different types of modulation and number of data packets transmitted at each time slot.

The key challenge of this model is to find the optimal time slot vector π_t to minimize the overall energy cost. By utilizing Equation 3.6 concerning the energy consumption via cellular networks, the optimal policies can be obtained by solving the following constrained optimization problem,

$$\min_{s_t \in S} \left\{ \gamma \cdot T_{delay} + \sum_{t=\tau_0}^{\tau_{k+m}} (f(d_{BS}, t) + E_{storage}(t)) \right\} \tag{5.3}$$

subject to:

$$\sum_{t=\tau_k}^{\tau_{k+m}} R_{tx}^t \cdot \tilde{t}(p_{pb}) \cdot \pi_t = F, \tag{5.4}$$

$$\pi_t \in \{0,1\}, \tag{5.5}$$

$$p^* = \arg\max E[\tilde{t}(p_{pb})], \tag{5.6}$$

Where γ is a constant denoting the importance (i.e., weight) of time delay compared with the energy consumption in transmission, R_{tx}^t is the transmission rate at time slot t. In this model, the data is restricted to be transmitted to the BS before a (hard) time deadline T. $\tilde{t}$ is the optimal time length to transmit the data to the BS in each time slot. The first step is using Equation 5.6 to calculate the optimal time duration in each time slot and distribution probability vector p^*. We assume that the traffic load of the PUs is unchanged within a short time duration (e.g. 60 seconds). Substituting Equation 3.10 and Equation 3.11 into Equation 3.9, we can obtain the optimal data transmission time duration in each time slot for the SUs. Due to the constant traffic load of the PUs in all of the channels, the probability of the SUs for transmission should be equal in each time slot.

Based on the above discussion, and utilizing the results from Equation 5.3, it is observed that the energy-efficiency optimization can in essence be formulated as an optimal stopping problem, which is to choose a time duration to minimize an expected cost [140]. The stopping decision would be made based on channel conditions, delay constraints, and the other factors of energy consumption which have been detailed above. To be specific, firstly, the policy calculates the data rate (bit/s) at each time slot and find the optimal time duration could be utilized for the data transmission of SUs in this time slot, thereby fixing the data length to be transmitted in this time slot. Secondly, this policy calculates the energy cost at all of the stopping time slots from the beginning spot along the route of vehicle, determining the last time slot (start from T) as the deadline for the data transmission. For instance, here is the procedure of calculating the first value of energy cost. At time slot τ_0, based on the data rate and optimal transmission time duration, the policy calculates the data length that would be transmitted, the cost for this transmission, and the remaining data size of the data. Repeat the process in the following time slots until the whole file has been transmitted to the BS completely. Finally, the policy should make sure that all of the data will be transmitted to the BS before the (hard) time deadline. The policy compared all the schemes which launch the data transmission at different time slots, finding the scheme with optimal trade-off between energy cost and time delay under deadline constraints.

5.3.1 OSP Formulation

The proposed OSP algorithm utilizes the information downloaded from navigation systems to fix the landmarks of nearby BSs and mobile users. Additionally, the model also takes advantage of the function of route prediction in order to trace the mobile terminals. Let S and D denote the starting point and terminus of mobile users. Let $Z_t^* = \{z_0, z_{\tau_1}, \cdots, z_{\tau_{k-1}}\}$ denote the set of user positions along the pre-computed route in each time slot, while Z_t represents the actual location of the mobile users updated by navigation systems. Under the time delay constraints T, Let ΔC_{pa} denote the parameter changes, and ε_{vp} and ε_{pa} denote the thresholds of distance of pre-computed vehicle

Algorithm 1 Energy cost optimization

Initialization:

Set $t = 0$; S; D; T; a set of location of nearby BSs; a set of parameters as described in Table 5.1 and Table 3.2;

Iteration:

while $(t < T \ \&\& \ F > 0)$ **do**
    if $\pi_t == 0$ **then**
        Break;
    else if $\pi_t == 1$ **then**
        transfer data from flash to DRAM;
        generate distribution probability vector p^*;
        RF module transmits data size l_t during current time slot;
        $F \leftarrow F - l_t$;
    end if
    if $|Z_t - Z_t^*| > \varepsilon_{vp}$ *or* $\sum \Delta C_{pa} > \varepsilon_{pa}$ **then**
        Recompute the set of π_t and Z_t^*;
        Update parameters;
    end if
    $t \leftarrow t + \Delta\tau$;
end while

Output:

optimal time slot vector π_t;

positions and parameters changes respectively. The energy cost optimization algorithm is illustrated in Algorithm 1 with optimal time slot vector π_t. $\Delta\tau$ is current time slot length and $\Delta\tau \in \Gamma = \{\tau_0, \tau_1, \tau_2, \cdots, \tau_j\}$. When mobile users are moving along the forecasted route, the algorithm must track the trajectory of the users and make sure that the parameters of vehicle, such as velocity, path, traffic lights, are not out of the system prediction. If any abnormal circumstance occurs, like changes of the predicted route, speed up or slow down abnormally, stop, traffic congestion or traffic accident, the system is required to have the capability to dynamically calculate the optimization of energy consumption again for the mobile devices and send the latest indicated message to notify the terminals.

5.3.2 Energy-Delay Trade-off

For highly elastic messages under different Delay constraints, we strive to select an optimal time interval for message transmission in order to achieve energy-efficiency with respect to delay requirements of mobile applications. Therefore, the energy-delay trade-off has emerged as a key concern aspect in cellular communications. We study the energy consumption for message transmission in a case that a vehicle is moving a certain distance towards the BS, which also takes into account the energy consumption of the DRAM and Flash embedded memory at the terminal which is used as a cost when message transmissions are delayed. The related values of the different parameters used in the numerical investigations are summarized in Table 3.2.

Similar to [1], we consider a four state channel with the following arrival rates: $\lambda_p^{(1)} = 0.01$, $\lambda_p^{(2)} = 0.01$, $\lambda_p^{(3)} = 0.02$, $\lambda_p^{(4)} = 0.02$, and $\lambda_s = 0.06$, the service time of PU: $E[X_p^{(1)}] = 20$, $E[X_p^{(2)}] = 30$, $E[X_p^{(3)}] = 20$, and $E[X_p^{(4)}] = 35$. As a SU, the wireless node on the vehicle would always try to select the optimal channel for data transmission and minimize the impact on the PU transmissions, i.e., the QoS of the current PU in the system should not be affected. From the above discussion, the optimal distribution probability vector for transmission time of vehicle p_{opt} is $(0.5308, 0.2828, 0.1864, 0)$, and the optimal transmission time duration in one time slot is $0.2354sec$, which have been specifically described in section 3.4. Our simulations assume that the traffic load is stable within a short time duration, which in this case is taken to be 50 seconds. In other words, the p_{opt} is constant in the whole process of message transmission.

Figure 5.3 and Figure 5.4 indicate the results of energy consumption starting from different time slots when a file of 50 Mbits is available for data transmission; in this case the vehicle initiated at a distance of 600 meters from the BS. The x-axis represents the launching message time slots as the vehicle is moving towards the BS, while the energy consumption is indicated on the y-axis. For example, the value of energy consumption at 20 sec at the x-axis presents the energy cost when the policy launches the message transmission at 20 sec time slot. As depicted in these two figures the results clearly indicate the inherent trade-offs between energy consumption and time delay of the proposed scheme over the CR networks. Figure 5.3 represents the energy consumption based on the system including RF, DRAM and NAND flash, and Figure 5.4 displays the energy curves with RF, DRAM and HDD. We can infer from the results that the energy cost will be significantly decreased if the system decides to transmit data after a short time delay. Furthermore, after a short period of time at the beginning, the proposed scheme of energy efficiency which transfers data between DRAM and NAND flash performs better than the one without storage transfer. At 20 sec, relative to the baseline model (without data transfer), the proposed scheme with data transfer ends up at saving $27.41J$, which

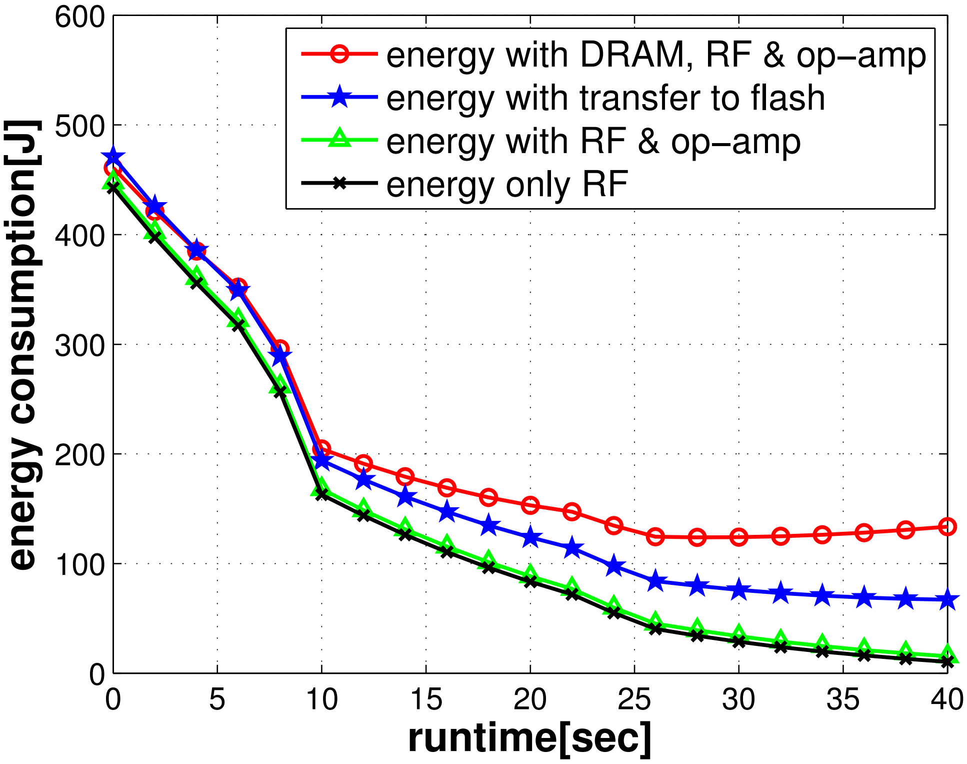


Figure 5.3: Energy consumption with DRAM, NAND flash from a distance of 600 meters to the BS

translates to an energy efficient benefit of 17.51%. The energy consumption represented by blue lines varies considerably, since compared with the power consumption of DRAM and NAND flash, the HDD has exceptionally higher power in the idle and read/write modes of operation.

When the vehicle is approaching closer to the serving BS, then compared to the RF transmission power consumption, power consumption of storage devices will become non-negligible in the prediction of transmission energy consumption as observed from the distinction between Figure 5.5A with the distance of 400 meters to the BS and Figure 5.5B where the distance is assumed to be 800 meters. Due to the short distance between wireless terminals and the BS, the energy consumption of storage devices will have an increasingly overall effect on the energy cost, which is vividly shown in the curves of Figure 5.5A.

By now, we have only set a time delay constraint on the transmission scheme protocol, but did not evaluate the importance of time delay for the whole transmission process. To

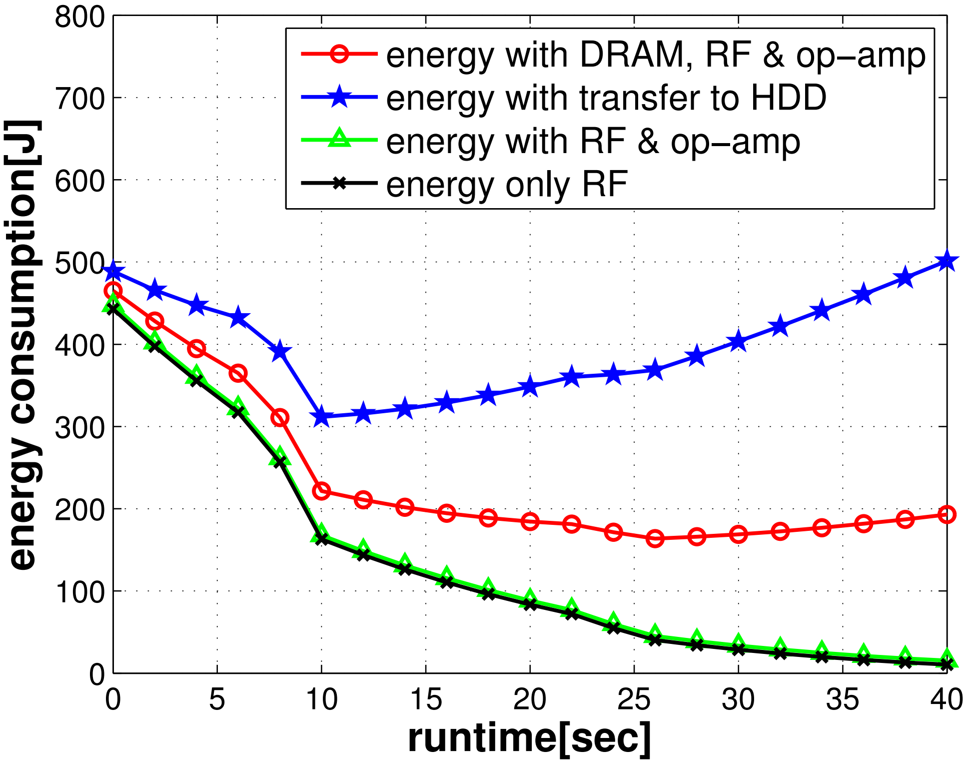
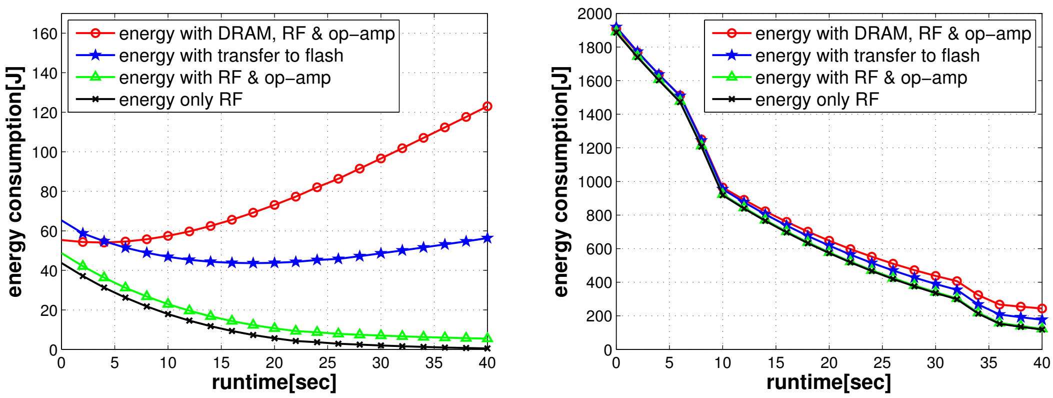


FIGURE 5.4: Energy consumption with DRAM, HDD from a distance of 600 meters to the BS



(A) 400 meters to Base Station

(B) 800 meters to Base Station

FIGURE 5.5: Energy cost from different distances to the BS

this end, different values of γ can be utilized to explore the importance of time delay in different applicable situations, such as the file transfer (delay-tolerant), video transmission (delay-sensitive). Figure 5.6 exhibits the difference of energy cost in different values of γ with 5, 10 and 20 sec time delay constraints. The hyaline part of the bars stands for the time delay incorporated into the overall energy consumption. $\gamma = 0$ reveals the

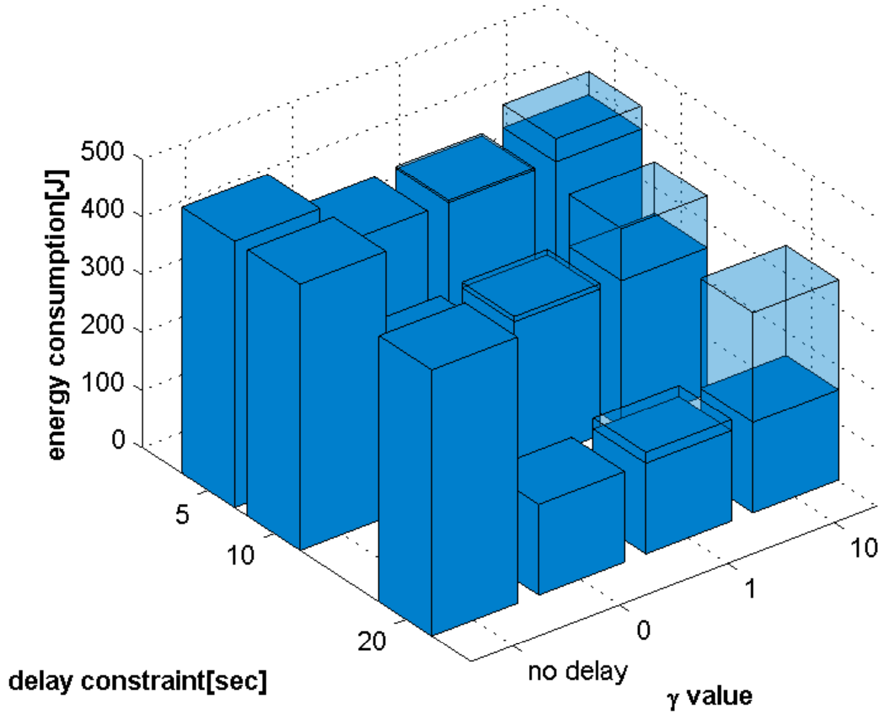


Figure 5.6: Importance of delay constraint in the overall energy cost

real value of energy consumption which delay cost has not been comprised. The case of $\gamma = 1$ and $\gamma = 10$ clearly demonstrate the rising importance of time delay in energy cost.

5.4 Transmission Cost VS Storage Cost

YouTube has become one of the most popular Internet applications, video clips are now streamed out not only to desktop computers but also to mobile devices. However, due to the constrained resource in mobile devices, YouTube like applications on portable devices could drain the battery within a short period. Once the mobile user make a demand to watch a YouTube video this will be streamed to the local DRAM of the mobile devices. When there are enough data stored in the DRAM then the applications will launch the video playback.

The statistical information regarding customer use could be utilized to explore the efficient use of battery in mobile devices. When a user chooses a specific video clip via the Internet, an HTTP message will be sent to the YouTube server to get the requested video. The multimedia server will respond and the content will be streamed out with the packets details, such as video identifier, size, source/destination IP addresses, and port numbers. In this manner, it is possible to determine the frequency of video request in unique digital devices. The research in [30] tracks the behavior of user requests from a campus network spanning an interval of 10 months. They use a commercial PC with a Data Acquisition and Generation (DAG) card to capture video information from YouTube server.

TABLE 5.2: YouTube Video Statistics per Digital Devices

Trace	Length(Hours)	Total Num	Single(%)	Multi(%)
1	12	12955	77.4	22.6
2	72	23515	77.7	22.3
3	108	17183	77.1	22.9
4	162	82132	72.5	27.5
5	336	303331	65.9	34.1
6	168	131450	68.5	31.5

The values in Table 5.2 from [30] show the statistics regarding the videos requested during the above mentioned track period. The 4th row (Single) presents the percentage of clips requested by one PC only once, while 5th row (Multi) shows the percentage of videos requested more than once. These statistics reveal that if the video could be buffered and cached in the local storages of digital devices, it will effectively decrease the wireless downlink traffic and energy consumption at the client end.

5.4.1 Energy Efficiency of Balancing Schemes

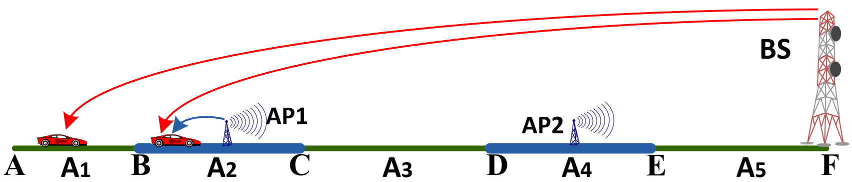


FIGURE 5.7: The selection from cellular BS and Wi-Fi AP

It is assumed that the road segment a vehicle is moving along is covered by a cellular BS with two Wi-Fi hotspots (segment *BC* and *DE*) in this domain as shown in Figure 5.7. We model the time duration until a popular video is requested again as an exponential distribution with mean time duration μ equal to 60 seconds; hence the rate at which content is requested is $\lambda = \frac{1}{\mu} = \frac{1}{60}$. Let $\mathcal{A} = \{A_1, A_2, \cdots, A_M\}$ represent the segments along the entire route when the vehicle is moving towards the BS and X denotes time interval the mobile user would potentially watch the same video. From the cumulative distribution function (CDF), the probability that the request takes place in the first segment (time duration $0 \sim t_1$) is $1 - e^{-t_1\lambda}$. Also, due to the memoryless property of exponential distribution, the probabilities in other time segments are given by $Pr(t < X < t + \Delta t \mid X > t) = 1 - e^{-\lambda \Delta t}$.

Let $\Gamma = \{\tau_i, \tau_{i+1}, \cdots, \tau_j\}$ denote the decomposed time slots in which mobile users is staying in a certain area from $\mathcal{A}$. Let $Q(t) = \{q_i, q_{i+1}, \cdots, q_j\}$ represent the energy consumed for wireless transmission in one area, such as electronic circuit at mobile devices, and energy consumed for data transmission. $E_{storage}(t)$ denotes the energy cost on storage devices. Our goal is to find the combined policies for overall energy efficiency in each area by minimizing the following expression:

$$E_{min} = \min_{s_t \in S(\mathcal{A})} \sum_{t=\tau_i}^{\tau_j} \left\{ Q(t) + E_{storage}(t) \right\} \tag{5.7}$$

Where s_t represents the strategy at time slot t, and S is the set of possible strategies including store and stream at each time slot in a certain area. E_{min} is the minimized energy cost for data storing and streaming in mobile users. Once a video clip has been downloaded from the multimedia servers and stored in the local DRAM already, there are two possible actions that could be taken. The first is that the content is deleted from the DRAM and hence future requests will have to be streamed again via wireless access. The other option is that the device stores the content in the local DRAM until a hard deadline. Our proposed set of schemes strive to balance between storing and transmission in the entire domain and explore the minimized energy cost across a long-term average, which is given as follows,

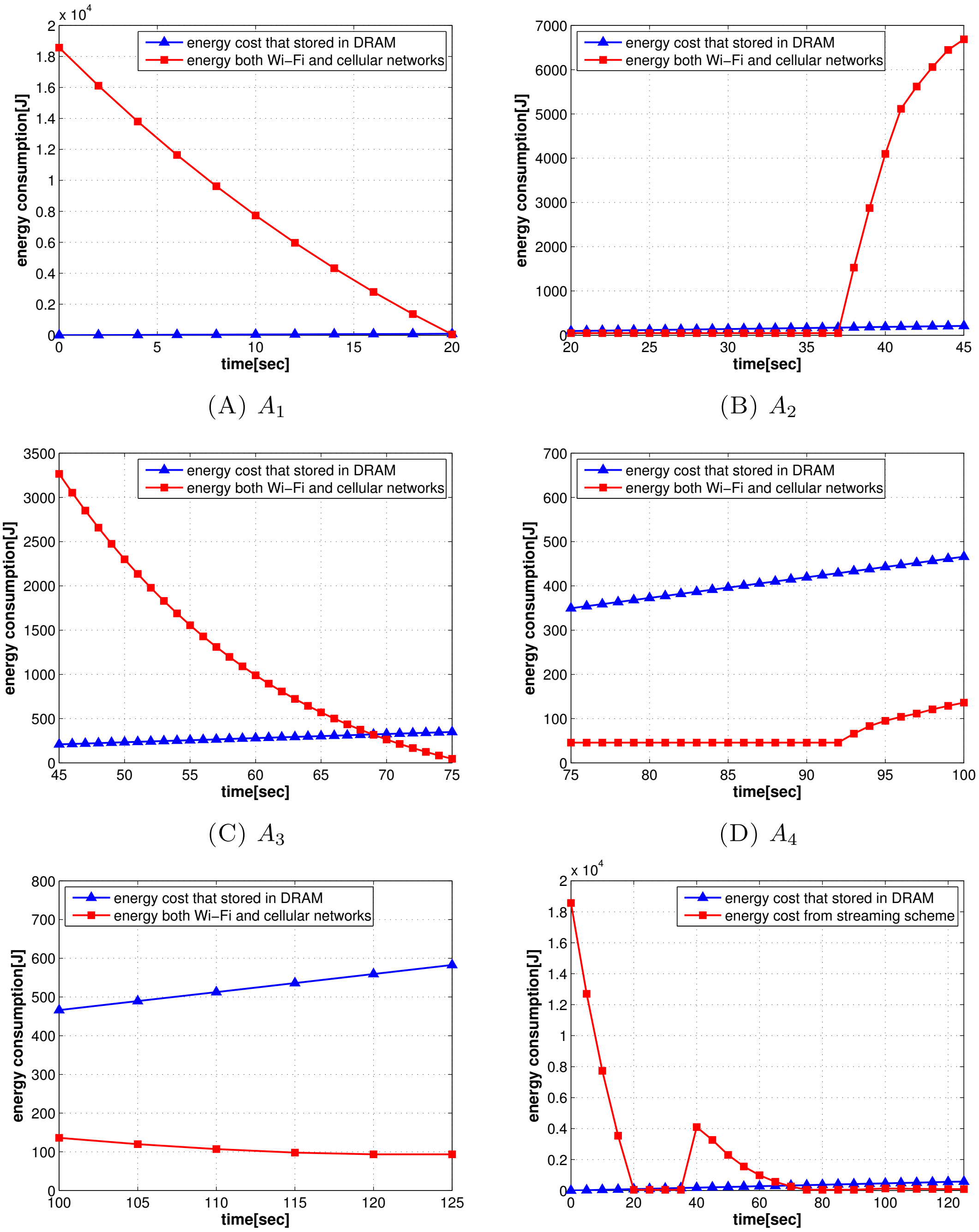
$$E_{opt} = \frac{\sum_{k=1}^{M} \{E_{min}(A_k) \cdot Pr(A_k)\}}{\sum_{k=1}^{M} Pr(A_k)} \tag{5.8}$$

Where $E_{min}(A_k)$ is the minimized energy cost in area A_k from the storing and streaming schemes as $A_k \in \mathcal{A} = \{A_1, A_2, \cdots, A_M\}$, while $Pr(A_k)$ is the corresponding probabilities that the mobile user would make a demand to consume the same content again. The proposed strategies combine transmission and selective storing aiming to achieve long-term energy efficiency according to the time interval distribution of user demand.

5.4.2 Numerical Investigations of Balancing Schemes

In this section numerical investigations are presented in several scenarios in which a vehicular user is moving towards a BS. Within the coverage of the BS, there are several Wi-Fi hotspots in which the user can connect to as passing by as shown in Figure 5.7. We assume the radius of the Wi-Fi hotspots to be 100 meters and the radius of the cellular network macro cell to be 1000 meters. Initially, the video clip has been downloaded and viewed from a multimedia server and stored in the local DRAM. As previously mentioned, the time interval that the mobile user would potentially watch the same popular video content again follows an exponential distribution. The video will be deleted from the local DRAM if the mobile user will not make a playing demand again before the threshold of 120 seconds. Therefore, there are 2 options for the mobile systems. 1) store the video clips in local DRAM until mobile users watch the stored video again; 2) delete the video from local DRAM immediately and always download via wireless access when mobile users make a request for this video. Table 3.2 presents the related parameters will be used in this section. Moreover, we assume the average multimedia video size is 10 MB on YouTube, which is a nominal value for such video files [129]. The results of energy consumption were obtained via MATLAB based simulation by focusing on the downlink throughput from BS and Wi-Fi AP. We also consider the coverage of AP in a cell of BS, and hence calculate the probabilities in different segments that a user may request the same content based on the previous mentioned exponential distribution.

In the first segment A_1, the probability that the mobile user would demand to watch the video again is $Pr(A_1) = Pr(X < 20) = 1 - e^{-\frac{20}{60}}$, and the related energy cost of two strategies is shown in Figure 5.8A. In the second segment A_2, the probability that the mobile user would require to watch the video again is given by $Pr(A_2|A_1^c) = Pr(20 < X < 45|X > 20) = 1 - e^{-\frac{25}{60}}$. The energy cost in this segment could be seen in Figure 5.8B. By this way, we could obtain the information regarding the energy cost in different areas as shown in Figure 5.8. The detailed information and figures concerning energy cost and corresponding probabilities in different areas are presented in Table 5.3.



(A) A_1 (B) A_2

(C) A_3 (D) A_4

(E) A_5 (F) entire process

FIGURE 5.8: Energy consumption along the road

Figure 5.8 demonstrates the energy cost on storing multimedia content in local DRAM and transmission cost in different areas. Figure 5.8A vividly demonstrates that the energy cost for wireless transmission is much higher than the DRAM cost due to the

Table 5.3: Probabilities and Energy Consumptions in Different Area

Area	energy for rx (Joule)	energy for store (Joule)	Probability (%)
A_1	45894.0	46.6088	0.2835
A_2	1509.8	151.4587	0.3408
A_3	2812.0	279.6088	0.3935
A_4	64.2	407.7587	0.3408
A_5	106.8	524.2587	0.3408

fact that the area under consideration is covered only by the BS which is located far away from the user. On the other hand, in the area A_2, the choice of preference is to stream the content from the Wi-Fi AP; note that however, there is a significant increase of the energy consumption for wireless transmission if the whole content cannot be transmitted completely during the time the user is within the Wi-Fi coverage area. In Figure 5.8C, there is an intersection point between store energy and transmission energy, from this time slot on, the energy cost for storing the multimedia content will outstrip the energy cost for wireless transmission. The figures in Table 5.3 suggest that in the area A_1, A_2, and A_3, the device should store the downloaded content in local DRAM in case the mobile user would request this content again in a short time period. Although in the area A_2 with Wi-Fi coverage, the majority of transmission energy is less than the energy to store it in the local DRAM, the energy consumption of storing the content in the local DRAM performs better than the schemes that always download via wireless access on a long-term average.

Let us consider the effect of this scheme on a long-term use. It is assumed that the case of 100 mobile users which are uniformly distributed within the coverage of the serving BS, and for all the users to request a popular video follows an exponential distribution with mean time 60 seconds. Figure 5.9 exhibits the energy cost of optimal schemes versus store/stream scheme on long-term average. By comparing energy cost for wireless transmission, when the mobile users require multiple times for the same video, it is a better strategy to store the video locally for a short period. Nevertheless, our combined schemes perform better among these cases based on the probabilities in different areas. The proposed scheme consumes 49.8% less energy compared to the storing scheme, while it is 5% more energy efficient from a scheme which is always streams the content. Finally, Figure 5.10 depicts the energy saving for an individual mobile user in one month, compared with the always streaming scheme with different

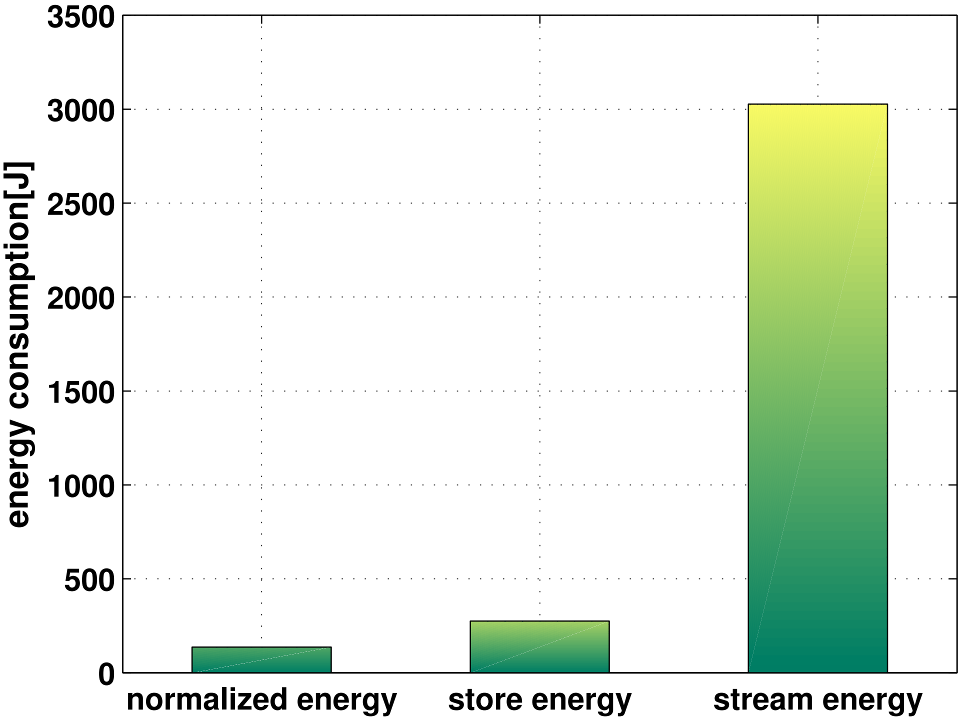


FIGURE 5.9: The average energy consumption of watching the same video for 100 users

elastic percentage and cumulative content size. These results suggest that as user request frequently the same video content, our proposed scheme could save a considerable portion of energy consumption, thus prolonging the lifetime of digital devices.

5.5 Cost Minimization

It is assumed that the road segment a vehicle is moving along is covered by a cellular BS with two White-Fi hotspots and two Wi-Fi hotspots in this domain as shown in Figure 5.11. Let $\Gamma = \{\tau_i, \tau_{i+1}, \cdots, \tau_j\}$ denote the decomposed time slots in which a mobile user is moving along a road. Let $Q(t) = \{q_i, q_{i+1}, \cdots, q_j\}$ represent the energy consumed by embedded peripherals, such as CPU, Graphics, backlight and storage devices. Our goal is to find the minimized energy cost for a fixed file size transmission over wireless access as follows,

$$E_{min} = \min_{s_t \in S} \sum_{t=\tau_i}^{\tau_j} \left\{ E_{cell}(t) + E_{TV}(t) + E_{WF}(t) + Q(t) \right\} \tag{5.9}$$

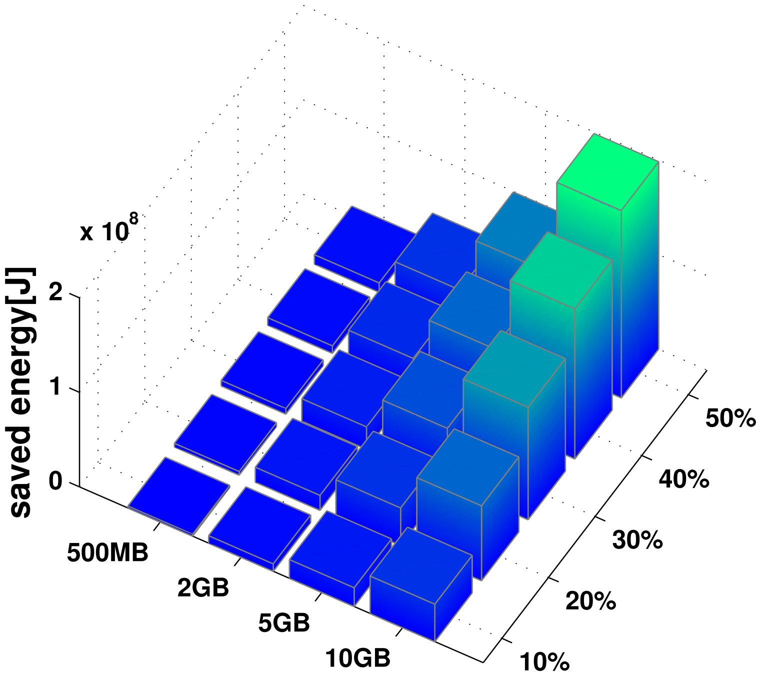


FIGURE 5.10: Saved energy consumption of the proposed over one month compared to the always streaming scheme

Where s_t represents the strategy at time slot t, and S is the set of possible strategies for wireless transmission. E_{TV} and E_{WF} represent the energy consumption for White-Fi and Wi-Fi components of mobile devices respectively. E_{min} is the overall energy cost for the embedded system in the mobile user end. Our proposed strategies strive to achieve energy efficiency for mobile users across a long-term average, thereby prolonging the battery lifetime.

5.5.1 Theoretical Analysis

In this section numerical investigations are presented in several scenarios in which a vehicular user is moving round downtown. Within the coverage of the BS, there are several White-Fi and Wi-Fi hotspots in which the user can connect to as passing by. We assume the radius of the White-Fi hotspots and Wi-Fi hotspots to be 100 and 50 meters respectively and the radius of the cellular network macro cell to be 1000 meters. Note that we only use the exponential distribution as an example here. Indeed,

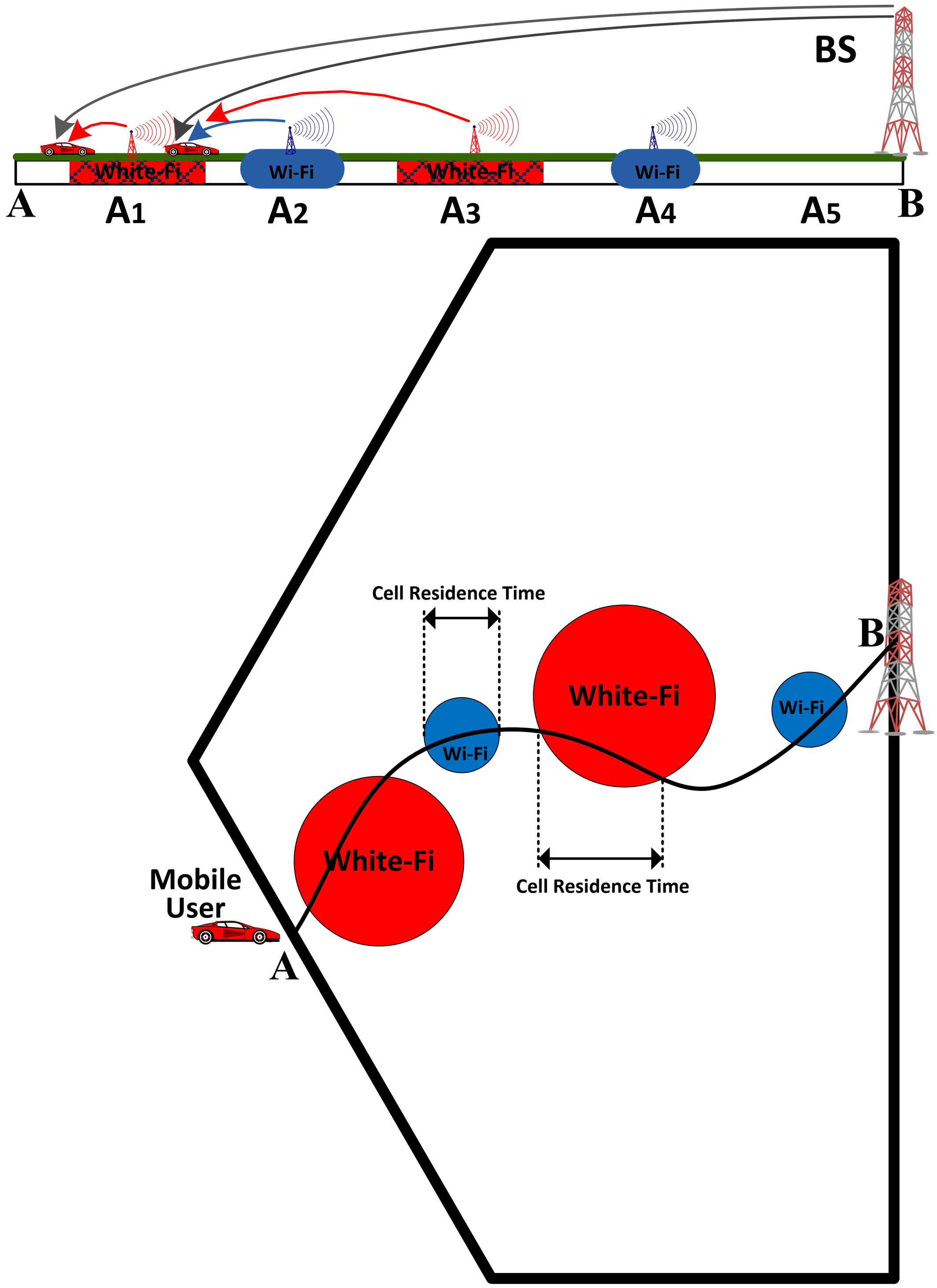


FIGURE 5.11: The selection from cellular BS, Wi-Fi and White-Fi AP

the proposed analytical framework can be applied to any distribution, e.g. gamma distribution, lognormal distribution. First, we simulate a simple network model with two White-Fi and Wi-Fi hotspots in order to show the results of our proposed schemes that can be clearly understood. The energy cost of main peripherals in mobile devices is based on the parameters from Table 4.3.

It is assumed that 125 mobile users are uniformly distributed within the coverage of the serving BS. Figure 5.12 exhibits the average energy cost of different file size with different delay constraints. Once the user makes a demand to mobile applications, according to the delay constraints of requirements, the system will plan the start point of wireless transmission to secure the applications' requirements on portable devices. If there already has been enough data of applications downloaded in the local storages of digital device, the mobile user will have more time to move into next high-capacity hotspot, as the interval would overrun the delay constraints. In Figure 5.12, we compare the energy cost of wireless transmission for different data size. Obviously, if the mobile applications allow for more delay, the device will have more time to move into a high-data rate hotspot, thus consuming less energy for wireless transmission. For Long Delay Tolerant Applications, we set the delay constraints for the mobile applications as $T_{delay} = 1min$, which could allow the mobile user moves into wireless hotspot area such as White-Fi and Wi-Fi. In this case, the wireless transmission can be always executed via the high throughput interfaces. In the case of Short Delay Tolerant Applications, we set the delay constraints for the mobile applications as $T_{delay} = 10sec$. Once the mobile user makes a demand to streaming video (YouTube like) applications on portable devices, the data will be streamed to the local DRAM of the mobile devices. If there are enough data stored in the DRAM for video play in one wireless hotspot, there would be more time for the mobile user to move into next high-capacity hotspot, which the interval would overrun the delay constraints.

5.5.2 Realistic Mobility Paths

We further examine how the proposed schemes perform under real network deployments and realistic user mobility patterns. To this end, we consider the area around Oxford Street in central London (UK) which can vividly present a representative actual traffic situation in a downtown of a big city. The locations of the APs (White-Fi, Wi-Fi) and BSs are presented in Figure 5.13[1]. The road network layout is imported from

[1]The locations of Base Stations considered in the simulations are extracted from http://www.sitefinder.ofcom.org.uk

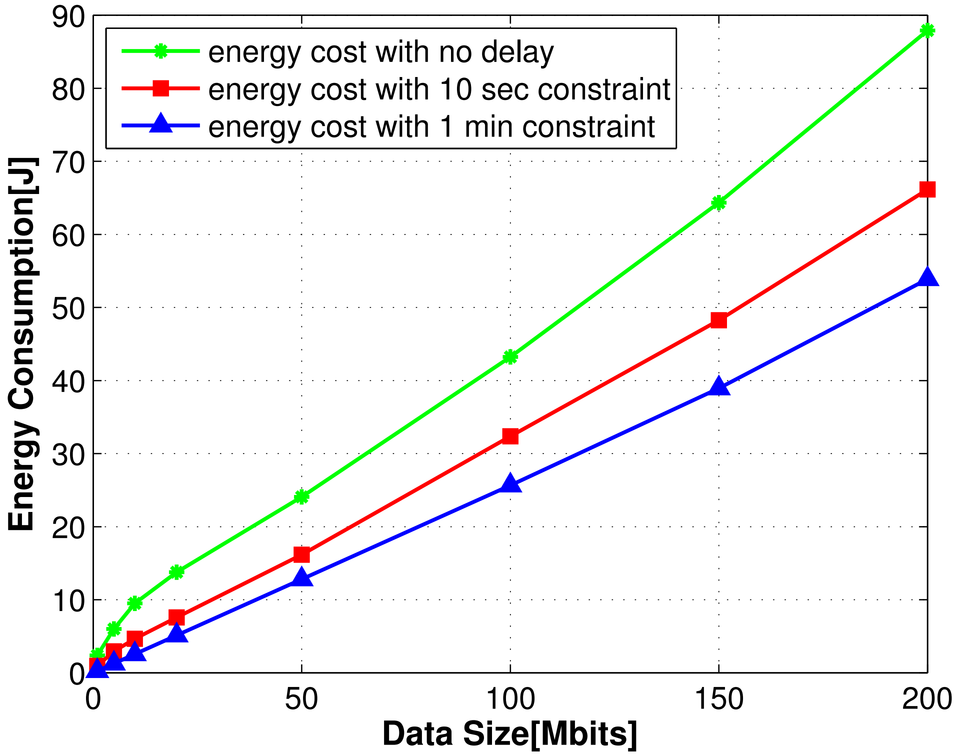


Figure 5.12: The average energy consumption of wireless transmission under different delay constraints

OpenStreetMap[2] into the SUMO[3] simulation tool (Appendix B) where continuous road traffic and realistic vehicular mobility of different types are being generated. Based on real-world networks, SUMO performs a traffic simulation consists of space-continuous and time-discrete vehicle movement, different vehicle types, multi-lane streets with lane changing, different right-of-way rules, and traffic light plans [141].

As already described in previous chapters, the average multimedia video size is approximately $10MB$ on YouTube [129]. For the assumed average transmission rates, the average length of such a video stream should be 4 minutes and 12 seconds, so in one cell, the mobile user will on average watch 250 seconds video which translates to about $100Mb$ downloading from wireless radios. In this deployment in central London, we select 10 different routes as shown in Figure 5.13 with 10 vehicles on each route from a cold start state, which means there is no traffic at the beginning of the simulations. For

[2]http://www.openstreetmap.org
[3]http://sumo-sim.org

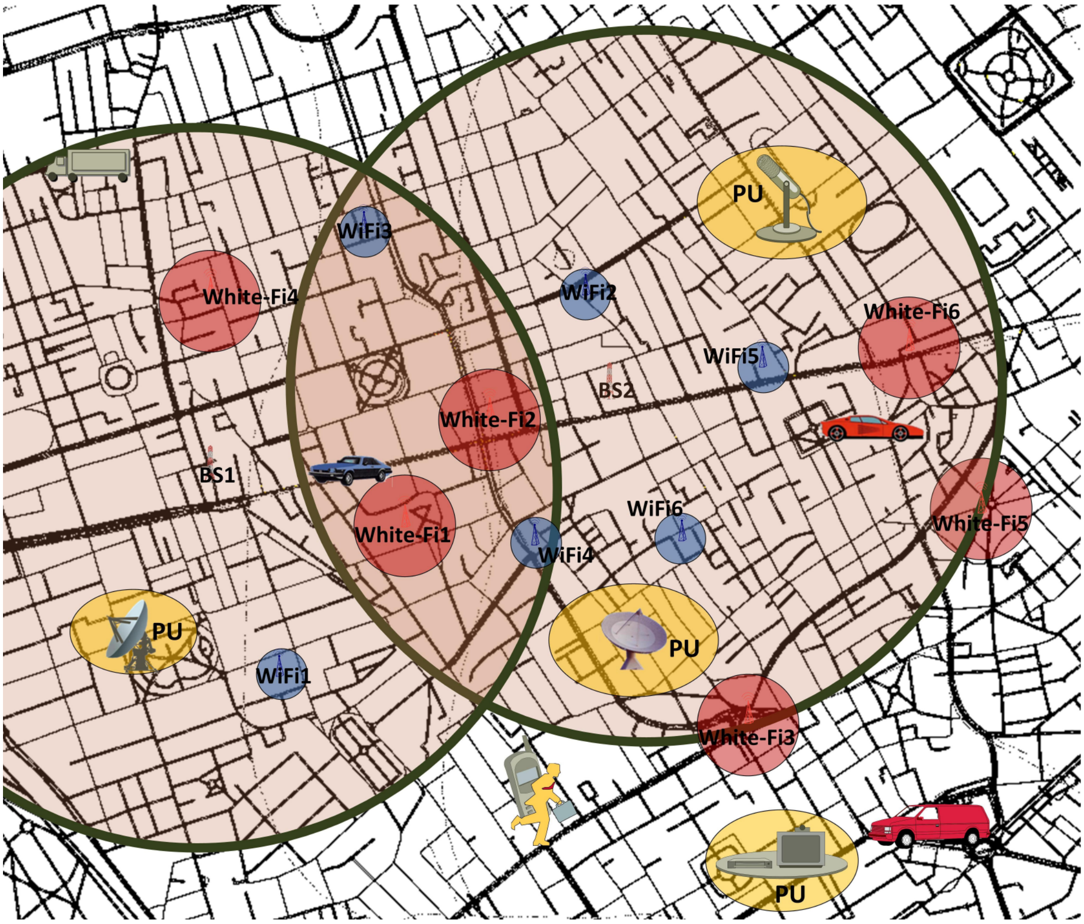
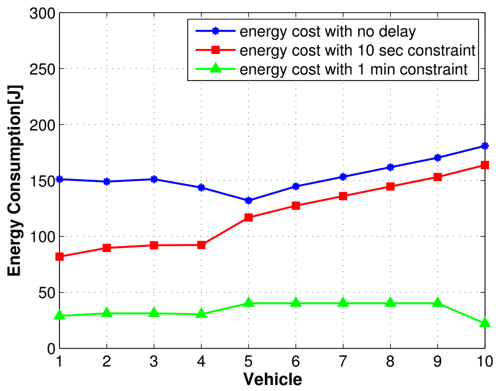


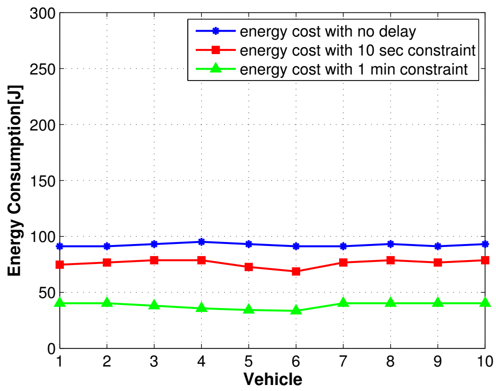
FIGURE 5.13: Topology for mobility simulation

every possible route each vehicle is moving continuously for 500 seconds and downloading $200Mb$ of data. If the delay tolerance of an application is high enough, in the orders of minutes for example, the most economic and energy efficient way of data downloading is via Wi-Fi or White-Fi. Similarly, the results of realistic mobility paths in Figure 5.14 show that the energy consumptions in three cases can be deemed as similar in the 10th path as shown in Figure 5.14J, across all 10 different mobility paths for delays up to 1 minute, the energy gains are on average 38.8%. Moreover, the minimum and maximum energy gains have been 19.1% and 82.9% respectively.

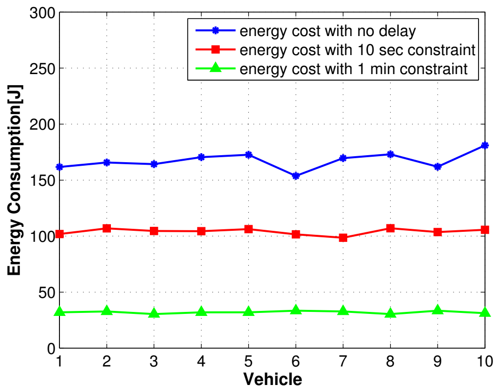
Figure 5.15 presents the energy cost of mobile devices in wireless transmission along the realistic mobility paths in a defined geographical area as shown in Figure 5.13. This graph depicts the variability of the energy consumption across different paths using different colours. Figure 5.15A displays the estimated energy cost in wireless transmission when there is no artificial delay included in the transmission, which means that



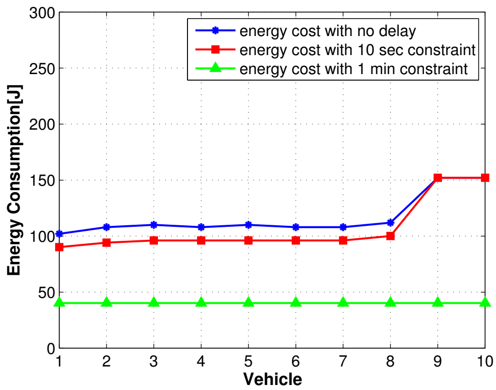
(A) 1st route



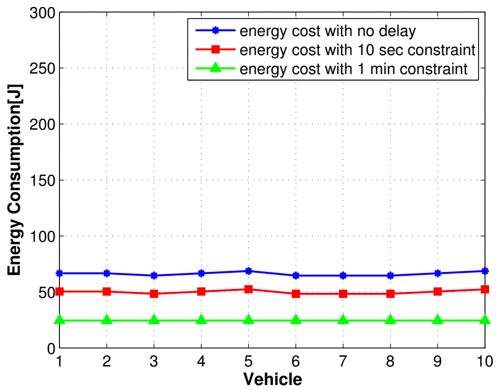
(B) 2nd route



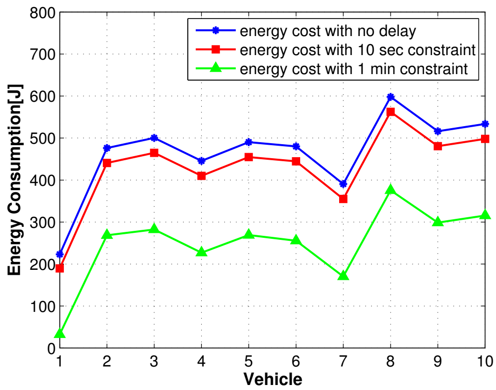
(C) 3rd route



(D) 4th route

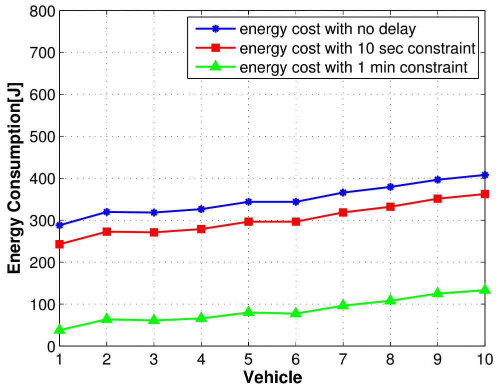


(E) 5th route

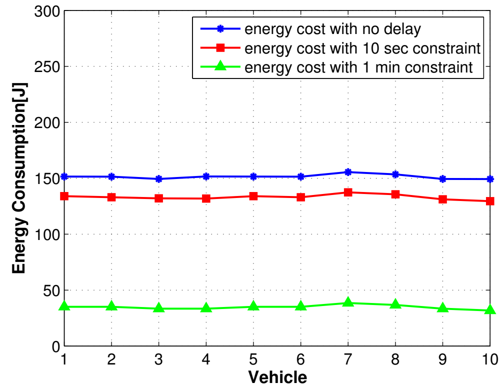


(F) 6th route

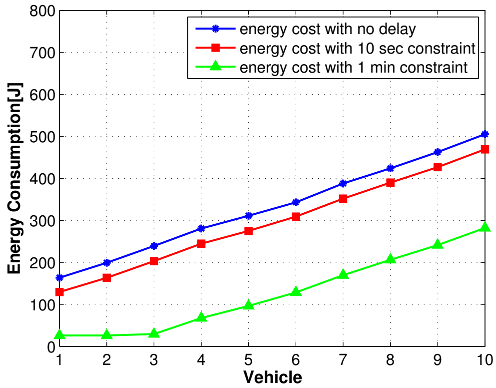
the mobile devices start data downloading immediately when requested by the mobile application. By comparing these two plots it can be clearly seen the spatial distribution of the performance improvement in the energy consumption when delays in the transmissions are allowed. For the case where delay is inhibited by the application, the mobility of the energy consumption area is coloured as red which means energy cost are more than 100J when mobile devices start wireless transmission at these points. The



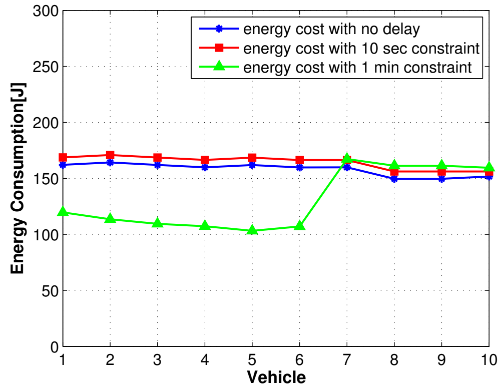
(G) 7th route



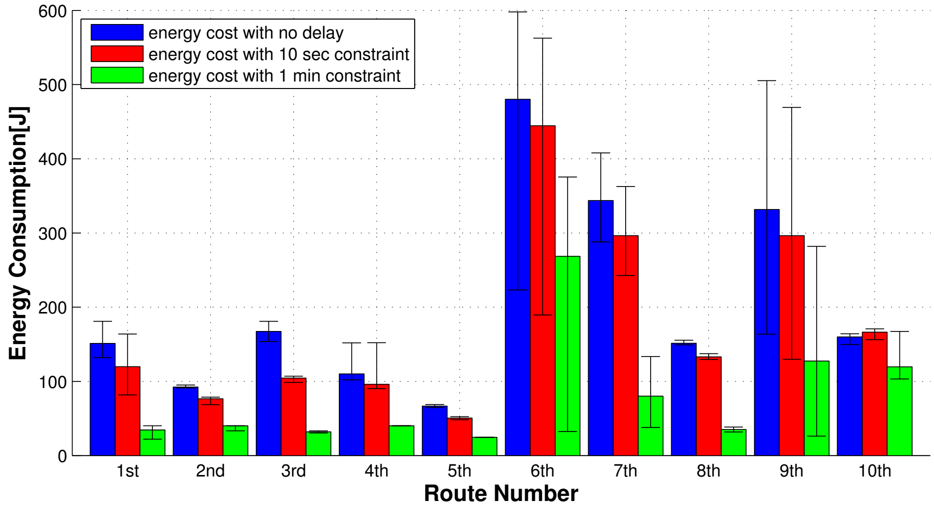
(H) 8th route



(I) 9th route

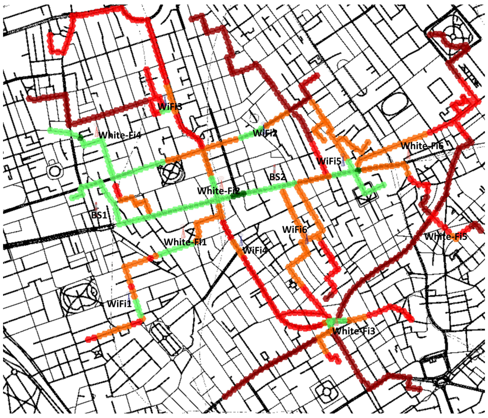


(J) 10th route

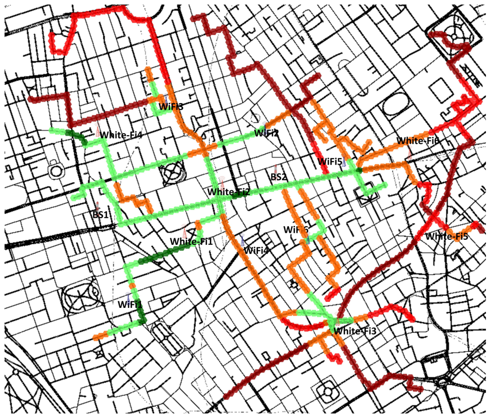


(K) Energy cost variation

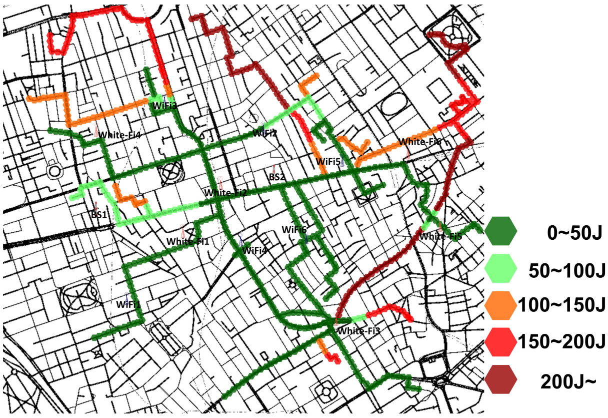
FIGURE 5.14: Energy cost from different routes



(A) no delay colored picture



(B) 10 Sec Delay



(C) 60 Sec Delay

FIGURE 5.15: Energy cost in different locations

mobile applications in Figure 5.15B allow the mobile systems have a 10-second delay; thereby it can be shown that some of pixels become green or orange since the reduction of energy cost in data download. Finally, if the applications' delay can be extended to 1 minute, a significant percentage of pixels are changing into dark or light green which translates to significant energy savings with respect to the former two cases as presented in Figure 5.15C. Note that in all the cases the same routes for mobile users apply. The scheme with 10-second delay offers on average gains of 86.9% compared to no delay scheme, while the scheme with 1 minute delay provides energy gains of 38.8% on average. The key observation of the above results is that due to the different delay thresholds for mobile applications, energy gains significantly increase when delay can be tolerated by the application, especially in areas which are populated with Wi-Fi, White-Fi cells. By default, the hotspots of Wi-Fi/White-Fi are inherently located in areas where mobile users will be more probably pass-by or where would be more traffic congestions that vehicles temporary stopping happened around there. Therefore, the use of these small cells can increase the performance as shown in this chapter.

5.6 Summary

Prolonging the recharging cycles of smartphones and mobile devices is considered as a prime objective as the proliferation of always-on Internet applications put significant strain on the battery capabilities. Previous research has revealed that the data transmission via wireless radios is a dominant energy consumption in mobile devices; thereby it is important to propose a new technique to avoid the drain of mobile batteries by intellectually utilizing the multiple radio interfaces in mobile terminals. When the SUs are using cellular networks in a geographic area, it is possible that the SUs would prefer the White-Fi or Wi-Fi connections over the cellular networks in terms of cost, RF coverage, capabilities, and transmission algorithm. Moreover, when the mobile user is relatively close to the BS, the energy consumption of the DRAM and flash memory in the embedded system can be of importance, which has been discussed in section 5.2.

On the other hand, the energy-delay trade-off has emerged as a key concern aspect in cellular communications. In section 5.3, a novel scheme for transmission scheduling has been explored in cognitive radio enabled networks for highly elastic messages, aiming to select best time interval for message transmission in order to achieve energy-efficiency under delay constraints. A key motivation for the proposed scheme is that when a vehicle is moving a certain distance towards the BS the energy consumption for wireless transmission declines significantly.

Finally, since energy consumption is a paramount metric for the usability of such portable devices, we study the trade-offs between storing popular video content locally at the DRAM of the device or allowing deleting the content from the local memory and relaying in wireless streaming in near-future requests of the same content. A scheme is proposed in section 5.4 where the mobility of the user is taken into account together with the probability of the user requesting the content multiple times so that a decision is taken of whether or not the content should be stored locally. Numerical investigations reveal that a combined scheme based on probabilistic analysis is essential for increasing the lifetime of digital devices, especially for the long-term energy efficiency on the wireless downlink transmission, which significant energy gains can be achieved in the order of 5% (from streaming scheme), and 50% (from storing scheme).

Chapter 6

Concluding Remarks and Future Directions for Research

In the final chapter, a synopsis of the thesis is presented in section 6.1, followed by a final conclusions gained from my research. Possible future directions for energy-efficient mobile systems on delay-tolerant applications are outlined in section 6.3.

6.1 Conclusions

Cisco estimated that by 2019 mobile video will account for nearly three-fourths of the world's mobile data traffic [2]. Within this aggregate traffic we envisage that the fraction of the mobile Internet video will steadily increase with the proliferation of smartphones and tablet devices. It is not only the significant increase in the mobile Internet traffic but a plethora of modern Internet applications are based on the always-on principle of operation and can inherently tolerate significant amounts of delays. In this environment, the energy limitations of smartphones have a direct impact on the length of time that mobile devices are operational. As previous investigations reveal data transmission via wireless radios is a dominant cost in mobile devices, and therefore it is important to understand the characteristics of wireless interfaces such as cellular networks, Wi-Fi and White-Fi. To this end, based on the accurate estimation of wireless networks condition, in Chapter 3, strategies are proposed to minimize the energy usage while meeting delay-tolerance deadlines specified by mobile users, which could intelligently make selective use of the high-rate wireless accesses. In subsection 3.1.3, an M/M/K/L queuing system is devised to estimate the probability for the SU connections and provides load-balancing of SU traffic by contacting a trusted database for historical information about PU traffic at a specific location and time duration. In the case of the delay-tolerant mobile applications, it is recommended to delay the data transmission to an area close to the serving BS in order to optimize the throughput potential while reducing the overall energy cost of wireless transmission subject to several constraints.

Stochastic characteristics of user mobility has been discussed in Chapter 4. Due to the fact that the overall cost of wireless transmission for mobile terminals can be greatly influenced by the time spent within the coverage of Wi-Fi/White-Fi Micro-cells, the mobility model of wireless terminals has been introduced in section 4.1. The characteristics of mobility pattern, such as speed changes of mobile users, direction changes, distributions of path selection and road restrictions, has been discussed in this section. Moreover, previous protocols for heterogeneous networks, such as MIPv6 and PMIPv6, have been introduced in section 4.2 to build a seamless relationship between the Macro-cell and the Micro-cells, which pursue efficient handoff latency and enhance vertical handover performance. Lastly, the mathematical models of cell residence time in wireless hotspots are studied in subsection 4.2.2 to calculate the file size transmitted via Micro-cells in order to better satisfy mobile application requirements, and optimize the switch cost of wireless radios in mobile systems.

Therefore, in subsection 4.2.3, based on a vehicle trajectory of mobile user within the coverage of Macro- and Micro- Cells in downtown, the transmission efficiency of two different schemes have been compared in subsection 4.2.3. The simulation results presented that the transmission performance of mobile terminals can be significantly enhanced by the vertical handover between the serving Macro-cells and Micro-cells, although the handover cost is introduced by switching among different radio interfaces. When the scenario is extended to long-term energy saving in section 4.3, if mobile applications are willing to tolerate some delays in exchange for high-quality transmission within the next available Wi-Fi/White-Fi hotspots, it is possible to prolong smartphone lifespan by up to 25%.

Meanwhile, we are currently witnessing the emergence of two important trends in wireless networks, the increased usage of delay-tolerant Internet-like applications and cognitive radio techniques. In Chapter 5, the focus is on how to capitalize the delay tolerance of various mobile applications to reduce the energy consumption in cognitive networks. According to the differences of delay tolerance for mobile applications, we group them into three categories; those who are delay sensitive (e.g., real time video meeting), those that can tolerate a short delay (e.g., video and audio streaming), and those who can stand significant long delay (e.g., RSS news feeds, email, software updates, etc.).

Based on the inherent mobility of wireless nodes, in section 5.3, an OSP methodology is introduced to seek optimal solutions for the wireless transmission scheduling by taking into account energy consumption and delay constraints together with available spectrum opportunities. Moreover, as part of energy consumption and in addition to various previous works, power consumption of storage devices have also been incorporated in section 5.2. It is found that it plays an important role in the overall energy cost (hence message delay has a cost in terms of Joules). This has been an issue that has not be previously taken explicitly into account, and the effect that different status of storage devices has been highlighted on the overall energy budget of mobile embedded systems.

As mobile users may want to watch the same popular video content multiple times, the trade-offs between wireless transmission cost and storage cost has been investigated in section 5.4. Once a video clip has been downloaded by mobile device, there are two possible actions that could be taken. One is storing popular video content locally at the DRAM of mobile device for a while, the other is allowing deleting the content from the local storages and relaying in wireless streaming in near-future requests of the same content. Therefore, a scheme combined transmission and selective storing is proposed to optimize long-term energy cost according to the time interval distribution of user demand. The experimental results have shown that, if mobile users can effectively avoid areas with higher energy cost for the elastic streaming media content, the proposed scheme could just consume up to 50% less energy compared to the worst case.

Finally, we introduced techniques that can reduce the energy cost for Internet applications with emphasis on the client side would effectively increase the battery lifetime of digital devices. To this end, it is necessary to introduce a strategy that deploys road-side hotspots and determines available spectrum at a given location to assist the data delivery for mobile devices according to corresponding on-line service features of mobile applications. Therefore, at last, in section 5.5, a practical model is proposed to optimize the overall energy cost of portable embedded systems for different mobile applications over wireless networks by incorporating and capitalizing the inherent mobility of the users. Experiments based on real world analysis using typical urban street environment reveal that impressive energy savings can be achieved for the delay-tolerant mobile flows which account for a significant portion of the total aggregate mobile Internet traffic. Once the mobile applications can tolerate more delay, by considering the distribution of the SU traffic load and PU connection that would be emerging stochastically, the mobile terminals would have more time to move into a high-speed Micro-cell, thus consuming considerably less energy for wireless transmission.

6.2 Publications

Book Chapters

- Bi Zhao, Vasilis Friderikos, "Towards Delay Tolerant Cognitive Cellular Networks", Green Communication (Wiley Book), 2013.

Journal Papers

- Bi Zhao, Vasilis Friderikos, "Increased Energy Efficiency via Delay-Tolerant Transmissions in Cognitive Radio Networks", Network Protocols and Algorithms, Vol. 5, No. 2, pp. 31-49, 2013.
- Bi Zhao, Vasilis Friderikos, "Utilizing Intermittent Small Cells Connectivity and User Mobility for Energy Efficiency in Delay Tolerant Applications", submitted.

Conference Papers

- Bi Zhao, Vasilis Friderikos, "Optimal Stopping for Energy Efficiency with Delay Constraints in Cognitive Radio Networks", PIMRC 2012, pp. 820-825, Sydney, Australia, Sept. 2012.

- Bi Zhao, Vasilis Friderikos, "A Queuing-Based Delay-Tolerant Scheme for Energy Efficiency over Cognitive Radio Networks", IEEE GLOBECOM Workshop on Emerging Technologies for Smart Devices (ETSD 2012), California, USA, Dec. 2012.

- Bi Zhao, Vasilis Friderikos, "Balancing Transmission and Storage Cost for Reducing Energy Consumption in Mobile Devices", VTC 2013-spring, Dresden, Germany, 2013.

- Bi Zhao, Vasilis Friderikos, "Extending Recharging Cycles of Mobile Devices with Intelligent Use of Wireless Interfaces", PIMRC 2013, London, UK, 2013.

6.3 Future Directions

The work presented in this thesis has provided different contributions within the DTNs and cognitive networks. The main stance to be defended is that by considering them jointly significant gains can be achieved in terms of energy savings at the terminal side. Although the experimental results show that impressive energy gain can be achieved over the delay-tolerant mobile applications and the battery lifetime of digital devices can be significantly extended, there are still some issues that can be improved in order to enhance the performance of Delay-tolerant Applications. Since cognitive users have to cease wireless transmission immediately and relocate to a new band as soon as a PU appears and requires access to the channel, SUs have to sense spectrum or query a database which maintains information about the available channels for the local radio environment as described in subsection 3.1.3. In the future, once the details of database regarding PU traffic is available for practical use, it is planned to study in a more detailed manner the technique that uses the available gaps in radio spectrum (the so-called 'white spaces'), which exist in the bands that have been reserved for analog TV broadcasting.

According to historical information concerning the movement of mobile users, it is important to predict the service time within the coverage of high-speed hotspots. As the differences of user actions, traffic conditions and road rules could lead to different cell residence time approximation by a particular distribution. In Chapter 4, only negative exponential distribution are employed to approximate the cell residence time for mobile users. Various distribution introduced in subsection 2.3.3 can be used, such as generalized gamma distribution, log-normal distribution, and truncated Gaussian distribution, etc., so as to find the distribution with optimal approximation, which depends on many factors such as cell shape, cell radius, user mobility and the selected paths.

The overall energy saving in mobile embedded systems has been evaluated in Chapter 5, which the aim is to select an optimized time duration to launch the wireless transmission. Hence the energy-efficiency optimization has been formulated as an optimal stopping problem, in which the storage devices in the embedded system can be of importance. As part of the future work in section 5.3, the challenge how to minimize the storage devices consumption by utilizing more detailed information from the application layer need to be investigated. Moreover, in section 5.4, different policies in terms of whether or not downloaded data should be stored in the terminals by taking into account the probability of re-using the data have been studied. From the report in [142], YouTube, Facebook and HTTP were the top 3 mobile downstream applications in North America and Europe. Since the downloaded data from the HTTP server could be held in a memory cache of mobile device [143], interesting future avenues of research would be to provide policies to store segments of content so that to provide better trade-offs between local storing and streaming the content via wireless transmission.

In section 5.5, the general model that a vehicle is moving along a route covered by a cellular BS and realistic path model are both considered. However, the speed of mobile users is another issue that should be considered for the time slots of data transmission. In the future, it is positive to discuss how the speed of the mobile users may affect the selection of time intervals for wireless transmission. Additionally, although the traffic congestion, traffic light, and road rule are considered in the case of realistic mobility, it would then be helpful to discuss the complexity of the proposed algorithms in section 5.5 and the sensitivity of the proposed to the dynamics related to user mobility, channel availability, and traffic conditions. Several features of the route diversity also need further investigation.

Appendix A

M/M/K/L Queuing System

This appendix describes the M/M/K/L queuing system utilized in Chapter 3. Consider a K-server queuing system with Poisson arrivals, exponential service times and finite number of waiting positions L. Clearly, total number of system places S = K + L. Assume that λ and μ are the arrival and service rates respectively as shown in Figure A.1. This queue is a variation of a multi-server system and only maximum (K + L) customers are allowed to stay in the system.

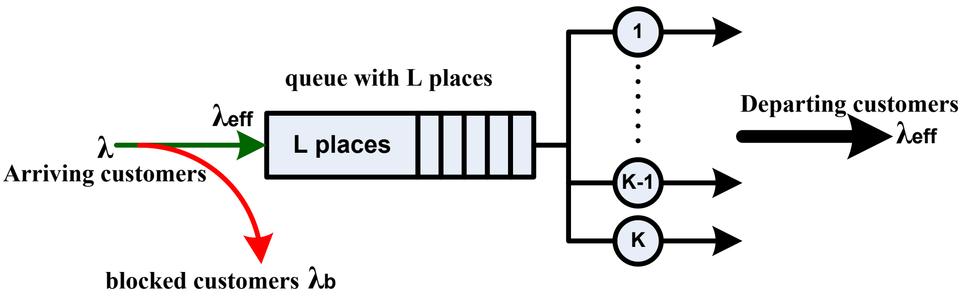


FIGURE A.1: M/M/K/L queuing system

Note that we assume the arrivals to be denied entry to the system (or the arrival process stops) once the number in the system reaches (K + L). The notation used in the M/M/K/L queuing system has been summarised in Table A.1. The system is in state m if there are m customers in the system (waiting or serviced), and let p_m be probability that there are m customers in the system in the steady-state, m = 0, 1, 2, $\cdots$, K+L. Therefore, according to the steady-state distribution we have

Table A.1: Notation Used in M/M/K/L Queuing System

Notation	**Description**
N	Average number of customers in the system, $N = N_q + N_s$
N_q	Average number of customers in the queue
N_s	Average number of customers in the service facilities
$\widetilde{x}$	Random variable which describes time spent in the service facility by a customer
$\overline{x}$	Average service time for a customer, $\overline{x} = E(\widetilde{x})$
$\widetilde{w}$	Random variable which describes time spent in the waiting queue by a customer
W	Average waiting time spent in the queue by a customer $W = E(\widetilde{w})$
$\widetilde{s}$	Random variable which describes time spent in the system by a customer; $\widetilde{s} = \widetilde{x} + \widetilde{w}$
T	Average time spent in the system by a customer $T = E(\widetilde{s}), T = W + \overline{x}$
λ	Arrival rate
λ_{eff}	The effective arrival rate
μ	Service rate
ρ	$\rho = \frac{\lambda}{\mu}$, offered load (offered traffic)
p_m	Stationary probabilities; p_m is the probability that there are m customers in the system

$$p_m = \begin{cases} \dfrac{\rho^m}{m!} \cdot p_0 & m \leq K \\ \\ \dfrac{\rho^K}{K!} (\dfrac{\rho}{K})^{m-K} \cdot p_0 & K < m \leq (K+L) \end{cases} \tag{A.1}$$

p_0 can be obtained using the condition $\displaystyle\sum_{m=0}^{K+L} p_m = 1$:

$$p_0 = \left[\sum_{m=0}^{K-1} \frac{\rho^m}{m!} + \sum_{m=K}^{K+L} \frac{\rho^m}{K! K^{m-K}} \right]^{-1} \tag{A.2}$$

Appendix B

SUMO (Simulation of Urban MObility)

This appendix introduces the SUMO simulation tool utilized in Chapter 5. SUMO is an open-source, microscopic, multi-modal traffic simulation for a given traffic demand, which consists of single vehicles moving through a given road network. It allows to address a large set of traffic management topics, in which each explicitly modelled vehicle has an own route, and moves individually through the traffic network. SUMO can be utilized to prepare and perform the simulation of a traffic scenario with following features:

- Space-continuous and time-discrete vehicle movement
- Different vehicle types
- Multi-lane streets with lane changing
- Different right-of-way rules, traffic lights
- Network Import includes: VISUM, Vissim, Shapefiles, OSM, RoboCup, MATsim, OpenDRIVE, and XML-Descriptions
- Microscopic routes - each vehicle has an own one
- High interoperability through usage of XML-data only

B.1 Networks/SUMO Road Networks

SUMO presents real-world road networks, where nodes (junctions) represent intersections, and roads are represented by edges. Intersections consist of a position, a shape, and right-of-way rules, which may be overwritten by a traffic light, while edges are unidirectional connections between two nodes and contain a fixed number of lanes with the following information:

- every street (edge) as a collection of lanes, including the position, shape and speed limit of every lane,
- traffic light logics referenced by junctions,
- junctions, including their right-of-way definitions; plain junctions first, then internal junctions
- connections between lanes at junctions (nodes)
- optionally roundabouts

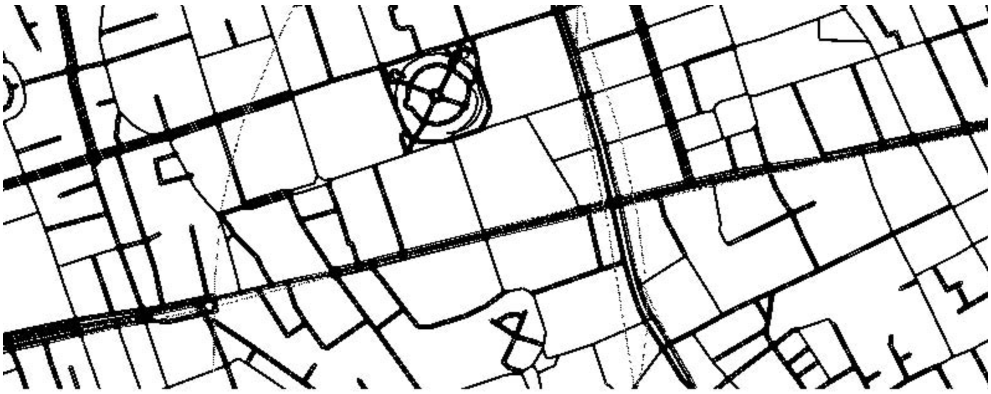


Figure B.1: A SUMO net file opened in SUMO-GUI

Figure B.1 displays a typical map of London downtown from a SUMO net file with filename extension ".net.xml" opened in SUMO-GUI. Although XML files are readable by human beings, a SUMO network file is not meant to be edited by hand. An existing map can either be converted from various formats of digital road map using NETCONVERT or generate geometrically simple, abstract road maps with NETGENERATE. Specifically, the road network importer NETCONVERT converts networks from other traffic simulators such as VISUM, Vissim, or MATSim. It also reads other common digital road network formats, such as shapefiles or OpenStreetMap.

B.2 SUMO Simulation

SUMO is a purely microscopic traffic simulation tool, in which each vehicle is explicitly defined at least by a unique identifier, the departure time, and the vehicle's route through the network. The route is presented by a complete list of connected edges between a vehicle's start and destination. If needed, each vehicle can be described in a finer detail using departure and arrival properties, such as the lane to use, the velocity, or the exact position on an edge. During the simulation, user can assign a set of variables for vehicles, which describe the vehicle's physical properties, movement model, and pollutant/noise emission classes with available options.

Meanwhile, SUMO performs a time-discrete simulation whereby the default step length is 1s, but can be lower down to 1ms. This is due to the fact that time is represented in microseconds stored as integer values. The simulation model is space-continuous. Internally each vehicle's position is described by the lane the vehicle is on and the distance from the beginning of this lane. When moving through the network, each vehicle's speed is computed using car-following models. Those models usually calculate an investigated vehicle's speed by factors such as this vehicle's speed, its distance to the leading vehicle, and the leader's speed.

Finally, SUMO can generate various outputs for each simulation, which include single vehicle positions in each time slots for all vehicles, information about each vehicle's trip, and aggregated measures for all streets and/or lanes. All output files are generated by SUMO in XML-format.

Bibliography

[1] Li-Chun Wang, Chung-Wei Wang, and F. Adachi. Load-Balancing Spectrum Decision for Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications*, 29(4):757–769, 2011.

[2] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014-2019. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html, February 2015.

[3] Lide Zhangy, Birjodh Tiwana, and et al. Accurate Online Power Estimation and Automatic Battery Behavior Based Power Model Generation for Smartphones. In *International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS 2010)*, pages 105–114, Scottsdale, AZ, October 2010.

[4] Eugene Shih, Paramvir Bahl, and Michael J. Sinclair. Wake on Wireless: An Event Driven Energy Saving Strategy for Battery Operated Devices. In *Proceedings of the 8th annual international conference on Mobile computing and networking (MobiCom 2002)*, pages 160–171, Atlanta, Georgia, USA, September 2002.

[5] G.P. Perrucci, F.H.P Fitzek, and J. Widmer. Survey on Energy Consumption Entities on the Smartphone Platform. In *Vehicular Technology Conference (VTC 2011 Spring)*, pages 1–6, Yokohama, Japan, May 2011.

[6] Trevor Pering, Yuvraj Agarwal, Rajesh Gupta, and Roy Want. CoolSpots: Reducing the Power Consumption of Wireless Mobile Devices with Multiple Radio Interfaces. In *Proceedings of the 4th international conference on Mobile systems, applications and services (MobiSys 2006)*, pages 220–232, Uppsala, Sweden, June 2006.

[7] Yuvraj Agarwal, Curt Schurgers, and Rajesh Gupta. Dynamic Power Management Using on Demand Paging for Networked Embedded Systems. In *Proceedings of Asia and South Pacific Design Automation Conference (ASP-DAC 2005)*, volume 2, pages 755–759, January 2005.

[8] Lei Lei, Zhangdui Zhong, Chuang Lin, and Xuemin Shen. Operator Controlled Device-to-device Communications in LTE-advanced Networks. *IEEE Wireless Communications*, 19(3):96–104, 2012.

[9] T.S. Rappaport, Sun Shu, R. Mayzus, Zhao Hang Zhao, Y. Azar, K. Wang, G.N. Wong, J.K. Schulz, M. Samimi, and F. Gutierrez. Millimeter Wave Mobile Communications for 5G Cellular: It Will Work! *IEEE Access*, 1(1):335–349, 2013.

[10] Wireless Gigabit Alliance. www.wirelessgigabitalliance.org, 2013.

[11] Karl Stetson. Wi-Fi Alliance and Wireless Gigabit Alliance finalize unification. http://www.wi-fi.org/news-events/newsroom/wi-fi-alliance%C2%AE-and-wireless-gigabit-alliance-finalize-unification, 2013.

[12] Sundeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges. *Proceedings of the IEEE*, 102 (3):366–385, 2014.

[13] Stephen J. Shellhammer, Ahmed K. Sadek, and Wenyi Zhang. Technical Challenges for Cognitive Radio in the TV White Space Spectrum. In *Information Theory and Applications Workshop*, pages 323–333, San Francisco, California, USA, February 2009.

[14] Ofcom. Digital Dividend: Cognitive Access, Statement on Licence-exempting Cognitive Devices Using Interleaved Spectrum. http://stakeholders.ofcom.org.uk/binaries/consultations/cognitive/statement/statement.pdf, July 2009.

[15] S. Probasco and Ed. B. Patil. Protocol to Access White Space database: PS, use cases and rqmts. http://tools.ietf.org/html/draft-ietf-paws-problem-stmt-usecases-rqmts-03, February 2012.

[16] Simon Haykin. Cognitive Radio: Brain-Empowered Wireless Communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, 2005.

[17] Ian F. Akyildiz, Won-Yeol Lee, Mehmet C. Vuran, and Shantidev Mohanty. NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: a Survey. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 50(13):2127–2159, 2006.

[18] Peter J. Smith, Abdulla Firag, Pawel A. Dmochowski, and Mansoor Shafi. Analysis of the M/M/N/N Queue with Two Types of Arrival Process: Applications to Future Mobile Radio Systems. *Journal of Applied Mathematics*, 2012.

[19] Niranjan Balasubramanian, Aruna Balasubramanian, and Arun Venkataramani. Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications. In *Proceedings of ACM/USENIX Internet Measurement Conference (IMC 2009)*, pages 280–293, Chicago, USA, November 2009.

[20] Trevor Pering, Vijay Raghunathan, and Roy Want. Exploiting Radio Hierarchies for Power-efficient Wireless Device Discovery and Connection Setup. In *18th International Conference on VLSI Design (VLSID05)*, pages 774–779, January 2005.

[21] Pablo Rodriguez, Rajiv Chakravorty, Julian Chesterfield, Ian Pratt, and Suman Banerjee. MAR: A Commuter Router Infrastructure for the Mobile Internet. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services (MobiSys 2004)*, pages 217–230, Boston, Massachusetts, USA, June 2004.

[22] Yuchen Wu, Yanmin Zhu, and Bo Li. Infrastructure-Assisted Routing in Vehicular Networks. In *Proceedings of the 31th conference on Information communications (INFOCOM 2012)*, pages 1485–1493, Orlando, Florida, March 2012.

[23] Ilias Leontiadis, Paolo Costa, and Cecilia Mascolo. Extending Access Point Connectivity through Opportunistic Routing in Vehicular Networks. In *Proceedings of the 29th conference on Information communications (INFOCOM 2010)*, pages 486–490, San Diego, California, USA, March 2010.

[24] Vladimir Bychkovsky, Bret Hull, Allen Miu, Hari Balakrishnan, and Samuel Madden. A Measurement Study of Vehicular Internet Access using in Situ Wi-Fi Networks. In *Proceedings of the 12th annual international conference on Mobile*

computing and networking (MobiCom 2006), pages 50–61, Los Angeles, California, USA, September 2006.

[25] Pralhad Deshpande, Anand Kashyap, Chul Sung, and Samir R. Das. Predictive Methods for Improved Vehicular WiFi Access. In *Proceedings of the 7th international conference on Mobile systems, applications, and services (MobiSys 2009)*, pages 263–276, Krakow, Poland, June 2009.

[26] Christopher Thraves, Gilberto A. Urueta, Pablo Vidales, and Marcin Solarski. Driving the Deployment of Citywide Ubiquitous WiFi Access. In *Proceedings of First International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SIMUTools)*, pages 1–8, 2008.

[27] Anthony J. Nicholson and Brian D. Noble. BreadCrumbs: Forecasting Mobile Connectivity. In *Proceedings of the 14th ACM international conference on Mobile computing and networking (MobiCom 2008)*, pages 46–57, San Francisco, California, USA, September 2008.

[28] Ahmad Rahmati and Lin Zhong. Context-for-Wireless: Context-Sensitive Energy-Efficient Wireless Data Transfer. In *Proceedings of the 5th international conference on Mobile systems, applications and services (MobiSys 2007)*, pages 165–178, San Juan, Puerto Rico, USA, June 2007.

[29] Moo-Ryong Ra, Jeongyeup Paek, and et al. Energy-delay Tradeoffs in Smartphone Applications. In *Proceedings of the 8th international conference on Mobile systems, applications, and services (MobiSys 2010)*, pages 255–270, San Francisco, California, USA, June 2010.

[30] Michael Zink, Kyoungwon Suh, Yu Gua, and Jim Kurose. Characteristics of YouTube network traffic at a campus network - Measurements, models, and implications. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 53(4):501–514, 2009.

[31] Andrew Lipsman. Major Mobile Milestones in May: Apps Now Drive Half of All Time Spent on Digital. http://www.comscore.com/Insights/Blog, June 2014.

[32] Xiukui Li and Seyed A. Zekavat. Traffic Pattern Prediction and Performance Investigation for Cognitive Radio Systems. In *Wireless Communications and Networking Conference (WCNC)*, pages 894–899, Las Vegas, NV, March 2008.

[33] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal. Energy-efficient Packet Transmission over A Wireless Link. *IEEE/ACM Transactions on Networking*, 10(4): 487–499, 2002.

[34] Wanshi Chen, Michael J. Neely, and Urbashi Mitra. Energy-Efficient Transmissions With Individual Packet Delay Constraints. *IEEE Transactions on Information Theory*, 54(5):2090–2109, 2008.

[35] Xiliang Zhong and Cheng-Zhong Xu. Delay-Constrained Energy-Efficient Wireless Packet Scheduling with QoS Guarantees. In *Global Telecommunications Conference 2005 (GLOBECOM 2005)*, volume 6, pages 3336–3340, St. Louis, MO, December 2005.

[36] Alvin Fu, Eytan Modiano, and John Tsitsiklis. Optimal Energy Allocation for Delay-Constrained Data Transmission over a Time-Varying Channel. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM 2003)*, volume 2, pages 1095–1105, April 2003.

[37] Hongseok Kim and Gustavo de Veciana. Leveraging Dynamic Spare Capacity in Wireless Systems to Conserve Mobile Terminals' Energy. *IEEE/ACM Transactions on Networking (TON)*, 18(3):802–815, 2010.

[38] Christopher M. Sadler and Margaret Martonosi. Data Compression Algorithms for Energy-Constrained Devices in Delay Tolerant Networks. In *Proceedings of the 4th international conference on Embedded networked sensor systems (SenSys 2006)*, pages 265–278, Boulder, Colorado, USA, November 2006.

[39] Bo Han, Pan Hui, V. S. Anil Kumar, Madhav V. Marathe, Guanhong Pei, and Aravind Srinivasan. Cellular Traffic Offloading through Opportunistic Communications: a Case Study. In *Proceedings of the 5th ACM workshop on Challenged networks (CHANTS 2010)*, pages 31–38, Chicago, Illinois, USA, September 2010.

[40] Xuejun Zhuo, Wei Gao, Guohong Cao, and Yiqi Dai. Win-Coupon: An Incentive Framework for 3G Traffic Offloading. In *Network Protocols (ICNP)*, pages 206–215, Vancouver, BC, October 2011.

[41] Parag Himanshu Dave and Himanshu Balachadra Dave. *Design and Analysis of Algorithms*. Pearson Education, 2009. ISBN 8177585959.

[42] Aruna Balasubramanian and Ratul Mahajan Arun Venkataramani. Augmenting Mobile 3G Using WiFi. In *Proceedings of the 8th international conference on Mobile systems, applications, and services (MobiSys 2010)*, pages 209–222, San Francisco, California, USA, June 2010.

[43] A. Finamore, M. Mellia, M. Munafo, R. Torres, and S. Rao. YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC 2011)*, pages 345–360, Berlin, Germany, November 2011.

[44] Dinan Gunawardena, Thomas Karagiannis, Alexandre Proutiere, and Milan Vojnovic. Characterizing Podcast Services: Publishing, Usage, and Dissemination. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference (IMC 2009)*, pages 209–222, Chicago, Illinois, USA, November 2009.

[45] Feng Qian, Zhaoguang Wang, Yudong Gao, Junxian Huang, and et al. Periodic Transfers in Mobile Applications: Network-wide Origin, Impact, and Optimization. In *Proceedings of the 21st international conference on World Wide Web (WWW 2012)*, pages 51–60, Lyon, France, April 2012.

[46] The Mobile Internet Report. http://www.morganstanley.com/institutional/techresearch/mobile_internet_report122009.html, 2009.

[47] Martin May, Gunnar Karlsson, and Vincent Lenders. A System Architecture for Delay-Tolerant Content Distribution. In *Proceedings of IEEE Conference on Wireless Rural and Emergency Communications (WreCom)*, pages 1–6, Rome, Italy, October 2007.

[48] Kevin Fall. A Delay-tolerant Network Architecture for Challenged Internets. In *Proceedings of the 2003 conference on Applications, technologies, architectures,*

and protocols for computer communications (SIGCOMM 2003), pages 27–34, Karlsruhe, Germany, August 2003.

[49] Rerngvit Yanggratoke, Abdullah Azfar, Mara Jos Peroza Marval, and Sharjeel Ahmed. Delay Tolerant Network on Android Phones: Implementation Issues and Performance Measurements. *Journal of Communications*, 6(6):477–484, 2011.

[50] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (IMC 2007)*, pages 1–14, California, USA, October 2007.

[51] Phillipa Gill, Martin Arlittz, Zongpeng Li, and Anirban Mahantix. YouTube Traffic Characterization: A View From the Edge. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (IMC 2007)*, pages 15–28, California, USA, October 2007.

[52] Qianchuan Zhao, Stefan Geirhofer, Lang Tong, and Brian M. Sadler. Opportunistic Spectrum Access via Periodic Channel Sensing. *IEEE Transactions on Signal Processing*, 56(2):785–796, 2008.

[53] Gosan Noh, Jemin Lee, and Daesik Hong. Stochastic Multichannel Sensing for Cognitive Radio Systems: Optimal Channel Selection for Sensing with Interference Constraints. In *Vehicular Technology Conference Fall (VTC 2009-Fall)*, pages 1–5, Anchorage, AK, USA, September 2009.

[54] Junghyun Heo, Jungchae Shin, Jihee Nam, Yutae Lee, Joon Goo Park, and Ho-Shin Cho. Mathematical Analysis of Secondary User Traffic in Cognitive Radio System. In *Vehicular Technology Conference 2008 (VTC 2008-Fall)*, pages 1–5, Calgary, Alberta, Canada, September 2008.

[55] Shanshan Wang, Junshan Zhang, and Lang Tong. A Characterization of Delay Performance of Cognitive Medium Access. *IEEE Transactions on Wireless Communications*, 11(2):800–809, 2012.

[56] Hsien-Po Shiang and Mihaela van der Schaar. Queuing-Based Dynamic Channel Selection for Heterogeneous Multimedia Applications Over Cognitive Radio Networks. *IEEE Transactions on Multimedia*, 10(5):896–909, 2008.

[57] Simon Fischer, Petri Mahonen, Marcel Schongens, and Berthold Vocking. Load Balancing for Dynamic Spectrum Assignment with Local Information for Secondary Users. In *New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2008)*, pages 1–9, Chicago, USA, October 2008.

[58] Huahui Wang, Jian Ren, and Tongtong Li. Resource Allocation with Load Balancing for Cognitive Radio Networks. In *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pages 1–5, Miami, FL, December 2010.

[59] Ho Ting Cheng and Weihua Zhuang. Simple Channel Sensing Order in Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications*, 29(4):676–688, 2011.

[60] Rahul Urgaonkar and Michael J. Neely. Opportunistic Scheduling with Reliability Guarantees in Cognitive Radio Networks. *International Journal of Mobile Computing*, 8(6):766–777, 2009.

[61] Ian F. Akyildiz, Won-Yeol Lee, Mehmet C. Vuran, and Shantidev Mohanty. A Survey on Spectrum Management in Cognitive Radio Networks. *IEEE Communications Magazine*, 46(4):40–48, 2008.

[62] Anh Tuan Hoang and Ying-Chang Liang. Adaptive Scheduling of Spectrum Sensing Periods in Cognitive Radio Networks. In *IEEE Global Telecommunications Conference 2007 (GLOBECOM 2007)*, pages 3128–3132, Washington, DC, November 2007.

[63] Osvaldo Simeone, Yeheskel Bar-Ness, and Umberto Spagnolini. Stable Throughput of Cognitive Radios With and Without Relaying Capability. *IEEE Transactions on Communications*, 55(12):2351–2360, 2007.

[64] Chung-Wei Wang and Li-Chun Wang. Analysis of Reactive Spectrum Handoff in Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications*, 30(10):2016–2028, 2012.

[65] Chung-Wei Wang and Li-Chun Wang. Modeling and Analysis for Proactive-decision Spectrum Handoff in Cognitive Radio Networks. In *IEEE international conference on Communications (ICC 2009)*, pages 2764–2769, Dresden, June 2009.

[66] Amine Laourine, Shiyao Chen, and Lang Tong. Queuing Analysis in Multichannel Cognitive Spectrum Access: A Large Deviation Approach. In *Proceedings of the 29th conference on Information communications (INFOCOM 2010)*, pages 1–9, San Diego, CA, March 2010.

[67] Caoxie Zhang, Xinbing Wang, and Jun Li. Cooperative Cognitive Radio with Priority Queueing Analysis. In *IEEE International Conference on Communications (ICC 2009)*, pages 1–5, Dresden, Germany, June 2009.

[68] Chittabrata Ghosh, Srikanth Pagadarai, Dharma P. Agrawal, and Alexander M. Wyglinski. Queueing Theory Representation and Modeling of Spectrum Occupancy Employing Radio Frequency Measurements. In *2009 IEEE 70th Vehicular Technology Conference Fall (VTC 2009-Fall)*, pages 1–5, Anchorage, AK, USA, September 2009.

[69] Oussama Habachi and Yezekael Hayel. Optimal Sensing Strategy for Opportunistic Secondary Users In a Cognitive Radio Network. In *Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems (MSWIM 2010)*, pages 343–350, Bodrum, Turkey, October 2010.

[70] Pero Latkoski, Jovan Karamacoski, and Liljana Gavrilovska. Availability Assessment of TVWS for Wi-Fi-like Secondary System: A Case Study. In *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM 2012)*, pages 196–201, Stockholm, June 2012.

[71] Qinghai Xiao, Yunzhou Li, Ming Zhao, Shidong Zhou, and Jing Wang. Opportunistic Channel Selection Approach under Collision Probability Constraint in Cognitive Radio Systems. *Computer Communications*, 32(18):1914–1922, 2009.

[72] Guangxiang Yuan, Ryan C. Grammenos, Yang Yang, and Wenbo Wang. Performance Analysis of Selective Opportunistic Spectrum Access With Traffic Prediction. *IEEE Transactions on Vehicular Technology*, 59(4):1949–1959, 2010.

[73] Johanna Vartiainen, Marko Hoyhtya, Janne Lehtomaki, and Timo Braysy. Priority Channel Selection Based on Detection History Database. In *the 5th International Conference on Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM)*, pages 1–5, Cannes, June 2010.

[74] Marko Hoyhtya, Sofie Pollin, and Aarne Mammela. Classification-based Predictive Channel Selection for Cognitive Radios. In *2010 IEEE International Conference on Communications (ICC)*, pages 1–6, Cape Town, South Africa, May 2010.

[75] Yang Song, Yuguang Fang, and Yanchao Zhang. Stochastic Channel Selection in Cognitive Radio Networks. In *IEEE Global Telecommunications Conference 2007 (GLOBECOM 2007)*, pages 4878–4882, Washington, DC, November 2007.

[76] Won-Yeol Lee and Ian. F. Akyildiz. Optimal Spectrum Sensing Framework for Cognitive Radio Networks. *IEEE Transactions on Wireless Communications*, 7 (10):3845–3857, 2008.

[77] Lei Yang, Lili Cao, and Haitao Zheng. Proactive Channel Access in Dynamic Spectrum Networks. *Physical Communication*, 1(2):103–111, 2008.

[78] Bechir Hamdaoui. Adaptive Spectrum Assessment for Opportunistic Access in Cognitive Radio Networks. *IEEE Transactions on Wireless Communications*, 8 (2):922–930, 2009.

[79] Volker Pauli, Juan Diego Naranjo, and Eiko Seidel. Heterogeneous LTE Networks and Inter-Cell Interference Coordination. In *Nomor Research GmbH*, pages 1–9, Munich, Germany, December 2010.

[80] Stefan Parkvall, Erik Dahlman, George Jongren, Sara Landstrom, and Lars Lindbom. Heterogeneous Network Deployments in LTE. Ericsson Review, February 2011.

[81] Hiroyuki Ishii, Yoshihisa Kishiyama, and Hideaki Takahashi. A Novel Architecture for LTE-B. In *IEEE Globecom Workshops: International Workshop on Emerging Technologies for LTE-Advanced and Beyond-4G*, pages 624–630, Anaheim, CA, December 2012.

[82] Christoph Paaschy and Gregory Detal. Exploring Mobile/WiFi Handover with Multipath TCP. In *Proceedings of the 2012 ACM SIGCOMM workshop on Cellular networks: operations, challenges, and future design (CellNet 2012)*, pages 31–36, Helsinki, Finland, August 2012.

[83] Setiyo Budiyanto, Muhammad Asvial, and Dadang Gunawan. Implementation of Genetic Zone Routing Protocol (GZRP) in 3G-WiFi Offload Multi Base Station.

In *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pages 1–5, Miami, FL, December 2010.

[84] Pralhad Deshpande, Xiaoxiao Hou, and Samir R. Das. Performance Comparison of 3G and Metro-Scale WiFi for Vehicular Network Access. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement (IMC 2010)*, pages 301–307, Melbourne, Australia, November 2010.

[85] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli. Integration of 802.11 and Third-Generation Wireless Data Networks. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies (INFOCOM 2003)*, volume 1, pages 503–512, Athens, March 2003.

[86] Hanane Fathi, Ramjee Prasad, and Shyam Chakraborty. Mobility Management for VoIP in 3G systems: Evaluation of Low-latency Handoff Schemes. *IEEE Wireless Communications*, 12(2):96–104, 2005.

[87] Maike Kuhnert, Thang Tran, and Christian Wietfeld. Cost-Effective and Feasible Handoff Application for Mobile Phones. In *IEEE Vehicular Technology Conference (VTC Fall 2012)*, pages 1–5, Quebec City, QC, September 2012.

[88] Younghyun Kim, Sangheon Pack, Chung Gu Kang, and Soonjun Park. An Enhanced Information Server for Seamless Vertical Handover in IEEE 802.21 MIH Networks. *Journal of Computer Networks: The International Journal of Computer and Telecommunications Networking*, 55(1):147–158, 2011.

[89] Stefano Busanelli, Marco Martalo, Gianluigi Ferrari, and Giovanni Spigoni. Vertical Handover between WiFi and UMTS Networks: Experimental Performance Analysis. *International Journal of Energy, Information and Communications*, 2 (1):75–96, 2011.

[90] Asanga Udugama, Muhammad Umer Iqbal, Umar Toseef, Carmelita Goerg, Changpeng Fan, and Morten Schlaeger. Evaluation of a Network based Mobility Management Protocol : PMIPv6. In *IEEE 69th Vehicular Technology Conference (VTC Spring 2009)*, pages 1–5, Barcelona, April 2009.

[91] Qinxue Sun, Bo Hu, Shanzhi Chen, and Jingjing Zhang. An Implementation of Vertical Handoff in Heterogeneous Networks. In *5th International Conference*

on Wireless Communications, Networking and Mobile Computing (WiCom 2009), pages 1–4, Beijing, China, September 2009.

[92] D. Johnson, C. Perkins, and J. Arkko. Mobility Support in IPv6. http://www.ietf.org/rfc/rfc3775.txt, June 2004.

[93] Giuliana Iapichino and Christian Bonnet. Experimental Evaluation of Proxy Mobile IPv6: an Implementation Perspective. In *IEEE Wireless Communications and Networking Conference (WCNC 2010)*, pages 1–5, Sydney, Australia, April 2010.

[94] Ki-Sik Kong, Wonjun Lee, Youn-Hee Han, and Myung-Ki Shin. Handover Latency Analysis of a Network-Based Localized Mobility Management Protocol. In *IEEE International Conference on Communications, 2008 (ICC 2008)*, pages 5838–5843, Beijing, China, May 2008.

[95] Christian M. Mueller and Oliver Blume. Network-based Mobility with Proxy Mobile IPv6. In *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, 2007 (PIMRC 2007)*, pages 1–5, Athens, September 2007.

[96] Ed. S. Gundavelli, K. Leung, V. Devarapalli K. Chowdhury, and B. Patil. IETF refc-5213: Proxy Mobile IPv6. http://tools.ietf.org/html/rfc5213, August 2008.

[97] Linoh A.Magagula, Olabisi E. Falowo, and H. Anthony Chan. Enhancing PMIPv6 for Better Handover Performance among HeterogeneousWireless Networks in aMicromobility Domain. *EURASIP Journal onWireless Communications and Networking*, 2010(1):1–13, 2010.

[98] Linoh A. Magagula, H. Anthony Chan, and Olabisi E. Falowo. PMIPv6-HC: Handover Mechanism for Reducing Handover Delay and Packet Loss in NGWN. In *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pages 1–5, Miami, FL, December 2010.

[99] M.S. Bargh, B. Hulsebosch, H. Eertink, G. Hedijenk, J. Idserda, J. Laganier, A. R. Prasad, and A. Zugenmaier. Reducing Handover Latency in Future IP-based Wireless Networks: Proxy Mobile IPv6 with Simultaneous Bindings. In *International*

Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2008), pages 1–10, Newport Beach, CA, June 2008.

[100] Mahmood M. Zonoozi and Prem Dassanayake. User Mobility Modeling and Characterization of Mobility Patterns. *IEEE Journal on Selected Areas in Communications*, 15(7):1239–1252, 1997.

[101] Farooq Khan and Djamal Zeghlache. Effect of Cell Residence Time Distribution on the Performance of Cellular Mobile Networks. In *IEEE 47th Vehicular Technology Conference (VTC 1997)*, volume 2, pages 949–953, Phoenix, AZ, May 1997.

[102] Yan Zhang and Boon-Hee Soong. The Effect of Handoff Dwell Time on the Mobile Network Performance. *Wireless Personal Communications*, 31(3-4):221–234, 2004.

[103] Yan Zhang and Boon-Hee Soong. Handoff Dwell Time Distribution Effect on Mobile Network Performance. *IEEE Transactions on Vehicular Technology*, 54 (4):1500–1508, 2005.

[104] Yuguang Fang, Imrich Chlamtac, and Yi-Bing Lin. Modeling PCS Networks Under General Call Holding Time and Cell Residence Time Distributions. *Journal of IEEE/ACM Transactions on Networking (TON)*, 5(6):893–906, 1997.

[105] Felipe A. Cruz-Prez and Arturo Seguin-Jimnez. On the Functional Relationship between Channel Holding Time and Cell Dwell Time in Mobile Cellular Networks. In *2004 IEEE Wireless Communications and Networking Conference (WCNC 2004)*, volume 4, pages 2387–2392, Atlanta, Ceorgia, USA, March 2004.

[106] Stamatis Kourtis and Rahim Tafazolli. Evaluation of Handover related Statistics and the Applicability of Mobility Modelling in Their Prediction. In *The 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2000)*, volume 1, pages 665–670, London, September 2000.

[107] Anum L. Enlil Corral-Ruiz, Andrs Rico-Pez, Felipe A. Cruz-Prez, and Genaro Hernndez-Valdez. Residence Time Distribution in Mobile Cellular Systems with Link Adaptation. In *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pages 1–6, Miami, FL, December 2010.

[108] Rainer Schoenen and Bernhard H. Walke. On PHY and MAC Performance of 3G-LTE in a Multi-hop Cellular Environment. In *Wireless Communications, Networking and Mobile Computing (WiCom 2007)*, pages 926–929, Shanghai, China, September 2007.

[109] Hannes Ekstrom, Anders Furuskar, Jonas Karlsson, Michael Meyer, Stefan Parkvall, Johan Torsner, and Mattias Wahlqvist. Technical solutions for the 3G long-term evolution. *IEEE Communications Magazine*, 44(3):38–45, 2006.

[110] P. Kolios, V. Friderikos, and K. Papadaki. Ultra Low Energy Store-carry and Forward Relaying within the Cell. In *IEEE Vehicular Technology Conference*, pages 1–5, Anchorage, AK, September 2009.

[111] Sophocles J. Orfanidis. *Electromagnetic Waves and Antennas*. www.ece.rutgers.edu/~orfanidi/ewa, 2008.

[112] Salvatore Chiaravalloti, Filip Idzikowski, and Lukasz Budzisz. Power Consumption of WLAN Network Elements. Technical Report TKN Technical Report Series TKN-11-002, Telecommunication Networks Group, Technical University Berlin, August 2011.

[113] Xinyu Zhang and Kang G. Shin. E-MiLi: Energy-Minimizing Idle Listening in Wireless Networks. In *Proceedings of the 17th annual international conference on Mobile computing and networking (MobiCom 2011)*, pages 205–216, Las Vegas, Nevada, USA, September 2011.

[114] H.Q. Nguyen, F. Baccelli, and D. Kofman. A Stochastic Geometry Analysis of Dense IEEE 802.11 Networks. In *26th IEEE International Conference on Computer Communications (INFOCOM 2007)*, pages 1199–1207, AK, May 2007.

[115] Jorg Ott and Dirk Kutscher. Drive-thru Internet: IEEE 802.11b for "Automobile" Users. In *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, volume 1, pages 362–373, March 2004.

[116] Lei Xie, Qun Li, Weizhen Mao, Jie Wu, and Daoxu Chen. Achieving Efficiency and Fairness for Association Control in Vehicular Networks. In *17th IEEE International Conference on Network Protocols (ICNP 2009)*, pages 324–333, Princeton, NJ, October 2009.

[117] Reshma Syeda and Vinod Namboodiri. Energy-Efficiency of Cooperative Sensing Schemes in Ad hoc WLAN Cognitive Radios. *Journal of Green Engineering - Special Issue on Cognitive Radio and the Application to Green Communications*, 2(1):347–375, 2012.

[118] William D. Ivancic, Wesley M. Eddy, Dave Stewart, Lloyd Wood, James Northam, and Chris Jackson. Experience with Delay-Tolerant Networking from orbit. *International Journal Satellite Communications Networking*, 28(5-6):335–351, 2010.

[119] Artemios G. Voyiatzis. A Survey of Delay- and Disruption-Tolerant Networking Applications. *Journal of Internet Engineering*, 5(1):331–344, 2012.

[120] Nikolaos Bezirgiannidis and Vassilis Tsaoussidis. Packet Size and DTN Transport Service: Evaluation on a DTN Testbed. In *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT 2010)*, pages 1198–1205, Moscow, October 2010.

[121] Forrest Warthman. Delay- and Disruption-Tolerant Networks (DTNs) A Tutorial. http://ipnsig.org/wp-content/uploads/2012/07/DTN_Tutorial_v2.04.pdf, July 2012.

[122] Christian Bettstetter and Hannes Hartensteinand Xavier Perez-Costa. Stochastic Properties of the Random Waypoint Mobility Model. *Journal on Wireless Networks*, 10(5):555–567, 2004.

[123] Wei jen Hsu, Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Ahmed Helmy. Modeling Time-Variant User Mobility in Wireless Mobile Networks. In *IEEE 26th International Conference on Computer Communications. (INFOCOM 2007)*, pages 758–766, Anchorage, AK, May 2007.

[124] Implementing Mobile IPv6 - Cisco. http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipv6/configuration/15-2mt/ipv6-15-2mt-book/ip6-mobile.pdf, 2012.

[125] Nickolay Amelichev, Mikhail Krinkin, and Kirill Krinkin. Behavior of Network Applications during Seamless 3G/WLAN Handover. In *Proceedings of the 11th international conference and 4th international conference on Smart spaces and next generation wired/wireless networking (NEW2AN'11/ruSMART'11)*, pages 575–586, St. Petersburg, Russia, August 2011.

[126] Kwan L. Yeung and Sanjiv Nanda. Channel Management in Microcell/Macrocell Cellular Radio Systems. *IEEE Transactions on Vehicular Technology*, 45(4):601–612, 1996.

[127] Aaron Carroll and Gernot Heiser. An Analysis of Power Consumption in a Smartphone. In *Proceedings of the 2010 USENIX Annual Technical Conference (USENIXATC'10)*, pages 21–21, Boston, MA, USA, June 2010.

[128] Howard David, Chris Fallin, Eugene Gorbatov, Ulf R. Hanebutte, and Onur MutluA. Memory Power Management via Dynamic Voltage/Frequency Scaling. In *Proceedings of the 8th ACM International Conference on Autonomic Computing (ICAC)*, pages 31–40, Karlsruhe, Germany, June 2011.

[129] Average Web Page Size Septuples Since 2003. http://www.websiteoptimization.com/speed/tweak/average-web-page/, May 2011.

[130] Joshua B. Fryman, Chad M. Huneycutt, and Hsien-Hsin S. Lee. Energy Efficient Network Memory for Ubiquitous Devices. *IEEE Micro*, 23(5):60–70, 2003.

[131] Krishna T. Malladiy, Frank A. Nothafty, Karthika Periyathambi, Benjamin C. Leez, Christos Kozyrakisy, and Mark Horowitzy. Towards Energy-proportional Datacenter Memory with Mobile DRAM. In *2012 39th Annual International Symposium on Computer Architecture (ISCA)*, pages 37–48, Portland, OR, USA, June 2012.

[132] Micron Technology Inc. 4Gb: x16, x32 Mobile LPDDR2 SDRAM S4 Features. http://www.micron.com/~/media/Documents/Products/Data%20Sheet/DRAM/Mobile%20DRAM/Low-Power%20DRAM/LPDDR2/U80M_4Gb_mobile_lpddr2_s4_sdram.pdf, 2013.

[133] Anthony Hylick, Ripduman Sohan, Andrew Rice, and Brian Jones. An Analysis of Hard Drive Energy Consumption. In *IEEE International Symposium on Modelling, Analysis and Simulation of Computers and Telecommunication Systems (MASCOTS)*, pages 103–112, Baltimore, MD, 2008.

[134] Laura M. Grupp, Adrian M. Caulfield, Joel Coburn, Steven Swanson, Eitan Yaakobi, Paul H. Siegel, and Jack K. Wolf. Characterizing Flash Memory: Anomalies, Observations, and Applications. In *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO 42)*, pages 24–33, New York, NY, USA, December 2009.

[135] Vidyabhushan Mohan, Sudhanva Gurumurthi, and Mircea R. Stan. FlashPower: A Detailed Power Model for NAND Flash Memory. In *Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, pages 502–507, Dresden, March 2010.

[136] Donald Molaro, Hannes Payer, and Damien Le Moal. Tempo: Disk Drive Power Consumption Characterization and Modeling. In *IEEE 13th International Symposium on Consumer Electronics (ISCE 2009)*, pages 246–250, Kyoto, Japan, May 2009.

[137] Song Liu, Karthik Pattabiraman, Thomas Moscibroda, and Benjamin G. Zorn. Flikker: Saving DRAM Refresh-power through Critical Data Partitioning. In *16th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, pages 213–224, Newport Beach, California, USA, March 2011.

[138] Micron Technology Inc. Mobile DRAM Power-Saving Features and Power Calculations. http://www.micron.com/~/media/Documents/Products/Technical%20Note/DRAM/tn4612.pdf, 2012.

[139] Jelena Trajkovic, Alexander V. Veidenbaum, and Arun Kejariwal. Improving SDRAM Access Energy Efficiency for Low-power Embedded Systems. *ACM Transactions on Embedded Computing Systems (TECS)*, 7(3):1–21, 2008.

[140] Thomas S. Ferguson. Optimal Stopping and Applications Available. http://www.math.ucla.edu/~tom/Stopping/Contents.html, 2008.

[141] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent Development and Applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3, 4):128–138, 2012.

[142] Sandvine. The Global Internet Phenomena Report: 2H 2014. https://www.sandvine.com/downloads/general/global-internet-phenomena/2014/2h-2014-global-internet-phenomena-report.pdf, September 2014.

[143] Xin Li, Mian Dong, Zhan Ma, and Felix Fernandes. GreenTube: Power Optimization for Mobile Video Streaming via Dynamic Cache Management. In *Proceedings of the 20th ACM international conference on Multimedia (MM 2012)*, pages 279–288, Nara, Japan, October 2012.